Numerical studies of QCD renormalons in high-order perturbative expansions

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)
der Fakultät für Physik
der Universität Regensburg

vorgelegt von
Clemens Bauer
aus Regensburg

November 2012

Die Arbeit wurde von Prof. Dr. Gunnar S. Bali angeleitet.

Das Promotionsgesuch wurde am 24. Oktober 2012 eingereicht.
Das Promotionskolloquium fand am 10. Januar 2013 statt.

Prüfungsausschuss:
Vorsitzender: Prof. Dr. J. Lupton
1. Gutachter: Prof. Dr. G. S. Bali
2. Gutachter: Prof. Dr. V. Braun
Weiterer Prüfer: Prof. Dr. J. Fabian

Contents

Introduction

What stands in the limelight of particle physics at present is the *Large Hadron Collider* (LHC) at CERN where experimentalists recently reported the observation of a new particle that could be the notorious *Higgs boson* [1]. The Higgs search is only the most famous exponent of a number of LHC experiments in quest of missing pieces or, quite to the contrary, deficiencies of what is known as the *Standard Model* of particle physics.

Its oldest building block is *Quantum Electrodynamics* (QED) as part of the electroweak theory. QED was developed more than 60 years ago and delivers an extremely accurate, relativistic description of all electromagnetic interactions. It gave prominence to the idea of calculating physical processes in *perturbation theory* with the help of *Feynman diagrams*, graphs which visualize and structure this approach. Its main concept is that a given quantity R can be expanded as a power series,

$$R = \sum_n r_n \alpha^{n+1}(Q)\,.$$

Here, the expansion parameter α reflects the interaction strength at a characteristic energy scale Q. Needless to say that this attempt only makes sense if α is sufficiently small. The contributions $r_n\alpha^{n+1}$ then decrease with growing perturbative order n and a truncation at finite n facilitates approximative solutions. QED lends itself to perturbative calculations as $\alpha_{\mathrm{em}}(Q) \sim \frac{1}{137}$ at low energies Q. This explains in part why QED is among the theories which have been confirmed with greatest accuracy so far. In the case of the electron's anomalous magnetic moment, for instance, the agreement between $\mathcal{O}(\alpha_{\mathrm{em}}^5)$ theory (including electroweak and strong contributions) and experiment surpasses astonishing nine significant digits [2].

Credit is due to Dyson [3] who was the first to scrutinize the perturbative approach. He cast doubt on the convergence of perturbative expansions for asymptotically large orders n. The fact that this typically is not the case is well accepted by now. The reason for the divergent nature of series expansions is the impossibility to strictly exclude from Feynman diagrams those energy regions Q' that obstruct a perturbative treatment. In QED, the coupling α_{em} grows with increasing energy scale Q' until the point where $\alpha_{\mathrm{em}}(Q')$ is no longer suited as an expansion parameter. Still, the starting point for the calculation was an expansion at a lower scale Q. The breakdown of perturbation theory is signaled by a rapid growth of the coefficients r_n to compensate for the smallness of $\alpha_e(Q)$. At best, the series expansion is *asymptotic*, it entails a finite ambiguity beyond which the series cannot

be determined. One particular divergence pattern is called *renormalon* and it will be the subject of this work.

The issue of renormalon divergence resurfaced in the 1970's with the rise of *Quantum Chromodynamics* (QCD) as the prime candidate for a theory of the strong interactions. A peculiarity of QCD is that the coupling constant $\alpha_s(Q)$ changes rapidly compared to QED. What is more, QCD is *asymptotically free*: the coupling constant grows as the energy scale Q decreases. Hence, the situation is just the opposite of QED and the problematic region now is the *infrared* (IR), the limit of low energies. As one can always think of arbitrarily soft corrections to a hard physical process – which is well-suited for perturbation theory otherwise – the infrared renormalon problem in QCD is omnipresent and all the more troublesome.

The failure of perturbation theory can be explained within a framework called *Operator Product Expansion* (OPE). In essence, the perturbative calculation of R is not the full story and must be complemented by non-perturbative power corrections. These are accompanied by universal *condensates* which parametrize in QCD the rich effects of its (low-energy) vacuum structure. Once the OPE combines the perturbative result with the non-perturbative corrections, the ambiguities cancel. Any physical observable R is then well-defined, as one would have expected from the very beginning.

Early studies regarding the asymptotic nature of perturbative expansions were restricted to purely theoretical considerations. Although widely accepted, a vigorous proof for renormalon existence in QCD is still pending. Apart from its fundamental interest, this question is of increasing practical importance: in recent years, several computations in perturbative QCD (e.g., [4, 5]), heavy quark physics in particular (e.g., [6–8]), have reached such a level of precision that they are potentially sensitive to renormalons. They may limit the accuracy with which the masses of bottom and top quarks near threshold can be determined.

Already decades ago, it became possible to actually probe the large-n limit by inspecting toy models far simpler than full-fledged quantum field theories such as QED or QCD. Toy model calculations were not only performed analytically [9], but also numerically [10] with the advent of the computer age.

In the mid-1990's, a method called *Numerical Stochastic Perturbation Theory* (NSPT) was proposed [11] that opened the way to large-order perturbative calculations – even in theories as complex as QCD and with the help of a computer. NSPT coincides with other, usually non-perturbative techniques in so far as it relies on the *lattice* technique [12], a discretization of spacetime.

Despite the enormous increase of computing power in the last decades, lattice QCD simulations in search for renormalons are still a challenging task. Moreover, conceptual difficulties arise on the lattice which have to be overcome before renormalons can be spotted. Failed or inconclusive NSPT attempts in recent years started to cast doubt on the conjectured existence of renormalons in QCD. This motivates us to reconsider the problem.

The outline of this thesis is as follows: in Chap. 1, we state the foundations of continuum QCD, its main features and take the step towards a lattice formulation. The following Chap. 2 is devoted to NSPT, the stochastic method it is based on and some technical aspects of its implementation. Then, Chap. 3 introduces the OPE and asymptotic perturbative expansions. It reviews the status of renormalons in various contexts, with a particular focus on two observables: the self-energy of an infinitely heavy quark and the elementary plaquette. Lattice quantities and techniques which are of general use in lattice simulations are presented in Chap. 4. We finally apply all preceeding considerations to our own renormalon study (Chap. 5) and conclude with a short summary in Chap. 6.

1 Quantum Chromodynamics

1.1 The QCD Lagrangian in the continuum

The development of *Quantum Chromodynamics* (QCD), the theory of the strong interactions, was stimulated by the remarkable success of *Quantum Electrodynamics* (QED). QED was the first relativistic quantum field theory that gave prominence to the principle of local gauge invariance. That is, the equations of motion remain unaffected if the spinor fields $\psi(x)$ describing spin-1/2 particles such as the electron transform as

$$\psi(x) \to e^{i\theta(x)}\psi(x)\,, \tag{1.1}$$

where $\theta(x)$ is a real-valued scalar function of a spacetime position x. In the language of group theory, Eq. (1.1) represents a U(1) gauge transformation. The construction of the QED Lagrangian satisfies from the very beginning both U(1) gauge symmetry and Poincaré symmetry. The latter embodies relativistic invariance, so QCD must also meet this criterion if it is meant to be a relativistic theory. The gauge symmetry of QCD, however, takes a more complicated form, since SU(3) is the gauge group under which the relevant fields, again spinors $\psi_f(x), \bar{\psi}_f(x)$ now describing quarks and antiquarks (distinguished by their *flavor* f), transform:

$$\begin{aligned} &\psi_f(x) \to \Lambda(x)\psi_f(x)\,, \\ &\bar{\psi}_f(x) \to \bar{\psi}_f(x)\Lambda^\dagger(x)\,, \\ &\Lambda(x) = e^{i\theta^a(x)t^a} \in \mathrm{SU}(3)\,, \qquad a = 1,\ldots,8\,. \end{aligned} \tag{1.2}$$

The $\theta^a(x)$ are real-valued spacetime functions as before. Here and below, repeated indices imply the usual Einstein sum convention. Novel are the *generators* t^a, with the following properties:

$$\begin{aligned} (t^a)^\dagger &= t^a\,, & \mathrm{Tr}\, t^a &= 0\,, \\ \mathrm{Tr}\left(t^a t^b\right) &= \frac{1}{2}\delta^{ab}\,, & \left[t^a, t^b\right] &= i f^{abc} t^c\,. \end{aligned} \tag{1.3}$$

Here, the f^{abc} are real and totally antisymmetric structure constants. In SU(3), they do not vanish and neither does the commutator $\left[t^a, t^b\right]$. At its heart, this is what makes QCD different from Abelian theories such as QED.

Noether's theorem states that each continuous symmetry of a physical system implies

a conserved quantity. The U(1) gauge symmetry is responsible for the conservation of electric charge in QED. By analogy, the conserved quantity in QCD due to SU(3) gauge symmetry is called *color charge*.[1] (Anti)quarks are (anti)color triplets – with each color component being a Dirac spinor – and live in what is called the *fundamental* representation of SU(3). Higher representations of SU(3) exist and the t^a are not uniquely defined. One particular choice, $t^a = \lambda^a/2$, is given by the set of 3×3 Gell-Mann matrices λ^a,

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
\lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},
\end{aligned} \tag{1.4}$$

and therefore acts naturally on the quarks in the fundamental representation. The t^a are called generators of the SU(3) Lie algebra because they span the complete space of traceless Hermitian 3×3 matrices, for any such matrix H can be represented as

$$H = \theta^a t^a \,, \tag{1.5}$$

with suitably chosen 8-component angle θ. An arbitrary group element $\Lambda \in \mathrm{SU}(3)$ can in turn be constructed as

$$\Lambda = e^{i\theta^a t^a} \,. \tag{1.6}$$

The reason for the 8 degrees of freedom can also be understood by looking at the fundamental representation of SU(3), made up by complex 3×3 matrices M with

$$M^\dagger M = \mathbb{1} \,, \qquad\qquad \det M = 1 \,. \tag{1.7}$$

The two constraints reduce the degrees of freedom from 18 to 8, which generalizes to $N^2 - 1$ degrees of freedom for arbitrary SU(N).

In addition to Eq. (1.3), the generators of SU(N) in the fundamental representation

[1] Asymptotic (matter) states in QCD turn out to be color neutral. For the group SU(N_c) with number of colors $N_c = 3$, this gives a vivid analogy with Young-Helmholtz theory: the additive superposition of the colors red, green and blue is perceived as white by humans.

yield

$$\sum_a t^a_{ij} t^a_{jk} = C_F \delta_{ik} \,, \qquad C_F = \frac{N^2 - 1}{2N} \,. \tag{1.8}$$

Here, i, j and k label matrix entries and C_F is the eigenvalue of the *quadratic Casimir* operator[2] $t^a t^a$ in the fundamental representation. We will see that, among the representations of higher dimension, the octet or *adjoint* representation is of special interest. Its generators T^a are 8×8 matrices which satisfy

$$\begin{aligned} T^a_{bc} &= -if^{abc} \,, & a, b, c &= 1, \ldots, 8 \,, \\ \mathrm{Tr}\left(T^a T^b\right) &= \sum_{i,j} f^{ija} f^{ijb} = C_A \delta^{ab} \,, & C_A &= N \,. \end{aligned} \tag{1.9}$$

Now, C_A marks the eigenvalue of the quadratic Casimir operator $T^a T^a$ in the adjoint representation. Both *color factors* C_F, C_A regularly appear in QCD calculations and in the case of SU(3) they read

$$C_F = \frac{4}{3} \,, \qquad C_A = 3 \,. \tag{1.10}$$

We return to the discussion of QCD and its protagonists, quarks and gluons. The flavors mentioned above refer to the six types of quarks we know of in nature ($n_f = 6$). They differ not only in their masses m_f (by up to several orders of magnitude [13]), but also in the electric charges they carry: *up* (u), *charm* (c) and *top* (t) with fractional charge $\frac{2}{3}e$ in units of the electron charge $-e$, and *down* (d), *strange* (s) and *bottom* (b) with negative charge $-\frac{1}{3}e$. Just as its electric charge dictates how a quark participates in electromagnetic interactions, the color charge mentioned above determines its behavior under strong interactions.

In analogy to QED and the photon, gauge bosons as the mediators of strong interactions arise naturally from the impossibility to introduce a kinetic term in the Lagrangian that only consists of quark fields and their derivatives and at the same time complies with local gauge invariance. As a consequence, one is forced to introduce the so-called covariant derivative

$$D_\mu = \partial_\mu - igA^a_\mu(x) t^a \equiv \partial_\mu - igA_\mu(x) \,, \tag{1.11}$$

with a new field A_μ called *gluon* whose coupling to the quark field has a relative strength g. Additionally, one demands that A_μ transforms under gauge rotations like

$$A_\mu(x) \longrightarrow \Lambda(x) A_\mu(x) \Lambda^\dagger(x) - \frac{i}{g} \Lambda(x) \partial_\mu \Lambda^\dagger(x) \,. \tag{1.12}$$

[2] Its characteristic is that it commutes with every generator in a given representation.

With the previous definitions, we can write down the fermionic part of the QCD Lagrangian,

$$\mathcal{L}_F(x) = \sum_{f=1}^{n_f} \bar{\psi}_f(x) \left(i\not{D} - m_f \right) \psi_f(x) \,, \tag{1.13}$$

where we have introduced $\not{D} = \gamma^\mu D_\mu$ with the known Dirac matrices γ_μ.

The gauge field $A_\mu(x)$ is a massless boson with unit spin. As is apparent from Eq. (1.11), it is an 8-component vector particle living in the octet or adjoint representation of the group. Gluons are not color neutral, they carry at the same time color and anticolor. Eight such color states are possible and can be related to the above Gell-Mann matrices.

If we want to account for gluon dynamics, we have to complement Eq. (1.13) by a kinetic term $\mathcal{L}_G$. Once more we borrow from QED and define the field strength tensor as the commutator of two covariant derivatives,

$$\begin{aligned} G_{\mu\nu}(x) =& \frac{i}{g} \left[D_\mu, D_\nu \right] = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) - ig \left[A_\mu(x), A_\nu(x) \right] = \\ =& G^a_{\mu\nu}(x) t^a = \left(\partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + g f^{abc} A^b_\mu(x) A^c_\nu(x) \right) t^a \,. \end{aligned} \tag{1.14}$$

Now we see that the non-Abelian nature of SU(3) (embodied by the non-vanishing f^{abc}) gives rise to an additional term in the field strength that is unknown in QED or other Abelian theories. It leads to three-gluon and four-gluon interactions, so gluons couple to each other. The picture is consistent: gluons interact with particles carrying color charges. As they carry color themselves, they are able to self-interact. The opposite holds for photons in QED.

We demand $\mathcal{L}_G$ to be of *mass dimension*[3] four and to respect gauge and Lorentz invariance. The only realization is

$$\mathcal{L}_G(x) = -\frac{1}{2} \mathrm{Tr} \left[G_{\mu\nu}(x) G^{\mu\nu}(x) \right] = -\frac{1}{4} G^a_{\mu\nu}(x) G^{a\mu\nu}(x) \,, \tag{1.15}$$

where the trace extends over color indices. We are now in the position to assemble the QCD Lagrangian:

$$\mathcal{L}_{\mathrm{QCD}}(x) = \mathcal{L}_F(x) + \mathcal{L}_G(x) = \sum_{f=1}^{n_f} \bar{\psi}_f(x) \left(i\not{D} - m_f \right) \psi_f(x) - \frac{1}{4} G^a_{\mu\nu}(x) G^{a\mu\nu}(x) \,. \tag{1.16}$$

1.2 The path integral method

Treating the (formerly) classical fields as quantum operators and postulating canonical commutation relations, *canonical quantization* leads from classical to quantum field theories. Another quantization method is the *path integral* formalism. In Euclidian space, it

[3] Throughout, we use $\hbar c = 197.3269718(44)$ MeV fm [13] and work in natural units.

draws from an analogy to statistical mechanics, in which expectation values are obtained as ensemble averages with according Boltzmann weights.

In the path integral approach to QCD, the expectation value of a generic observable $R\left[A,\psi,\bar{\psi}\right]$ is obtained as

$$\langle R\rangle = \frac{1}{Z}\int \mathcal{D}\left[A,\psi,\bar{\psi}\right] R\left[A,\psi,\bar{\psi}\right] e^{iS\left[A,\psi,\bar{\psi}\right]}\,, \tag{1.17}$$

where the action S and the partition function Z are defined by

$$S\left[A,\psi,\bar{\psi}\right] = \int d^4x\, \mathcal{L}\left[A,\psi,\bar{\psi}\right]\,, \tag{1.18}$$

$$Z = \int \mathcal{D}\left[A,\psi,\bar{\psi}\right] e^{iS\left[A,\psi,\bar{\psi}\right]}\,, \tag{1.19}$$

and $\mathcal{D}\left[A,\psi,\bar{\psi}\right]$ is a properly chosen integration measure. Wick rotating the time coordinate,

$$t = x_0 \longrightarrow -ix_4\,, \tag{1.20}$$

we switch to Euclidian spacetime $x^E = (\vec{x}, x_4)$ such that the path integral Eq. (1.17) now reads

$$\langle R\rangle = \frac{1}{Z^E}\int \mathcal{D}\left[A,\psi,\bar{\psi}\right] R\left[A,\psi,\bar{\psi}\right] e^{-S^E\left[A,\psi,\bar{\psi}\right]}\,. \tag{1.21}$$

For simplicity, consider pure gauge theory. Note that the change of metric replaced oscillatory by Gaussian weights: S_G^E is real and bounded from below, as can be seen from the definition of $\mathcal{L}_G$ [Eq. (1.16)] and the relative minus sign it acquires when translating to Euclidian space,

$$\mathcal{L}_G^E = -\mathcal{L}_G^M\,. \tag{1.22}$$

Therefore, paths in configuration space with least action have the largest impact on expectation values. As we will see later on, Eq. (1.21) opens the door to numerical simulations within lattice QCD.

1.3 Asymptotic freedom and confinement

For any kind of interaction, one of the fundamental questions is to determine the interaction strength as a function of distance or, likewise, the related energy (momentum) scale Q.[4] Based on the coupling g introduced above, one defines the *strong coupling constant* as

$$\alpha_s(Q) = \frac{g^2(Q)}{4\pi}\,. \tag{1.23}$$

[4] They are inversely proportional.

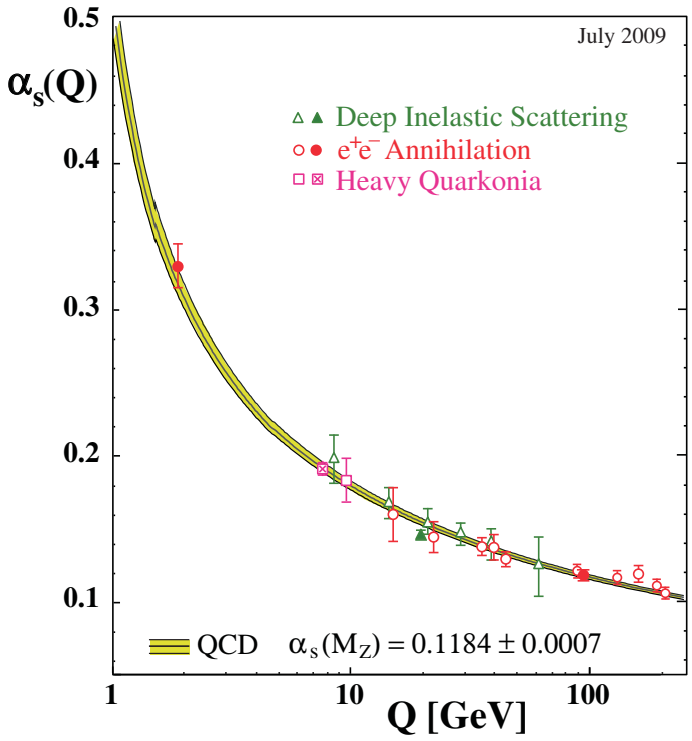

Fig. 1.1: The running coupling $\alpha_s(Q)$ as a function of the corresponding energy scale Q. Figure taken from [14].

We will drop the subscripts whenever it is clear that the strong interaction is implied. In the case of QED, the corresponding quantity $\alpha_{\text{em}}(Q) = e^2/(4\pi)$ is roughly constant indeed, with α_{em} ranging from $\sim 1/137$ at small energies $Q \approx 0$ up to $\sim 1/128$ at the scale of W boson mass m_W [13]. The reason for this slight increase is that an individual electric charge e is screened by a cloud of virtual electron-positron pairs. The closer one gets to the charge, the larger the measured value for e and hence α_{em} will be. QCD shows that this need not necessarily be the case: its *β-function*, which determines the running of α_s with scale μ (cf. Appendix A.1), yields

$$\beta(\alpha) = \mu \frac{d\alpha_s}{d\mu} = -\frac{\beta_0}{2\pi}\alpha_s^2 + \mathcal{O}\left(\alpha_s^3\right) . \tag{1.24}$$

One can separate gluonic and fermionic contributions, $\beta_0 = 11 - 2/3 n_f$. Since $n_f = 6$ for all that we know at present, $\beta_0 > 0$ and one sees that the non-Abelian structure of the gluon fields we encountered in Sec. 1.1 is responsible for the strong interaction's weakening with growing energy (cf. Fig. 1.1). At sufficiently high energies, the coupling becomes small enough to allow for perturbative calculations and one speaks of the *asymptotic freedom* [15, 16] that quarks experience, allowing for their isolated treatment in hard processes.

Yet towards lower energies one finds $\alpha_s \to \mathcal{O}(1)$, leading to the *confinement* of quarks inside color neutral matter states called *hadrons*. Perturbative calculations then become unfeasible, asking for an alternative method to address the rich and interesting physics of

low-energy QCD. *Lattice QCD* presents such an alternative and we will discuss its main features in the following.

1.4 Lattice QCD

Euclidian spacetime[5] can be discretized by means of a finite four-dimensional hypercubic lattice, with $N_1 \times N_2 \times N_3 \times N_4$ lattice nodes or *sites* in total. The smallest distance on the lattice is a, the separation between two neighboring sites. Each site is labeled by

$$x \to na \equiv (n_1, n_2, n_3, n_4)\, a\,, \qquad n_\mu \in \{0, 1, \ldots, N_\mu - 1\}\,. \tag{1.25}$$

Additionally, $\hat{\mu}$ shall denote a unit vector of length a in direction μ.

The finite lattice spacing a introduces a natural regulator on the lattice: the possible momentum components are bounded by π/a in the *ultraviolet* (UV). As a consequence, one avoids the singularities that can be incurred without cutoff. Additionally, finite computer resources require finite box extents $N_\mu a$. Its inverses put lower bounds on the momenta that can be reached in the *infrared* (IR). If only (bosonic) gauge fields are involved, the simplest way to deal with finite lattice extents is to introduce *periodic boundary conditions* (PBC)

$$(n_\mu + N_\mu)\, a \equiv n_\mu a\,. \tag{1.26}$$

The topology of the lattice then corresponds to a 4-dimensional torus. As we will see in Sec. 4.2, other boundary conditions are possible.

Having discretized spacetime, the next question is how to implement the quantum fields familiar from QCD on the lattice. We place the fermion fields $\psi, \bar{\psi}$ on the lattice sites, i.e.,

$$\psi(x) \to \psi(na)\,. \tag{1.27}$$

The gauge fields A_μ reside on the parallel transporters between adjacent sites n and $n+\hat{\mu}$. One possible definition of these so-called *links* is given by

$$U_\mu(x) \equiv e^{igA_\mu\left(n+\frac{\hat{\mu}}{2}\right)a}\,. \tag{1.28}$$

The same link passed in reverse direction then yields

$$U_{-\mu}(x+\hat{\mu}) = U_\mu^\dagger(x)\,. \tag{1.29}$$

Let us see how we can carry over the continuum QCD definitions of Chap. 1 to the lattice (for an illustration, see Fig. 1.2). The gauge condition for the gluon fields [Eq. (1.12)]

[5] From now on, the superscript E indicating the Euclidian metric is dropped but understood throughout.

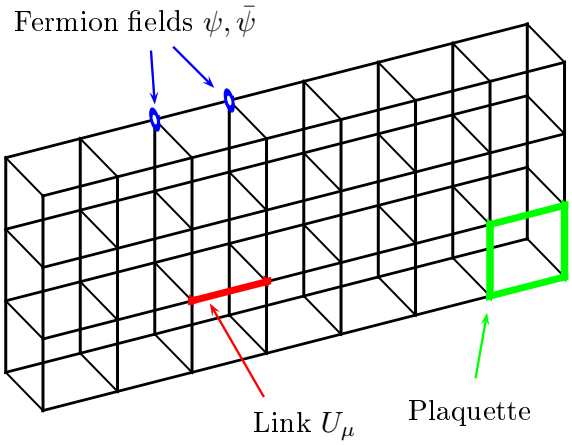

Fig. 1.2: Visualization of basic lattice objects.

translates into

$$U_\mu(x) \longrightarrow \Lambda(x) U_\mu(x) \Lambda^\dagger(x+\hat{\mu})\,, \tag{1.30a}$$

$$U_\mu^\dagger(x) \longrightarrow \Lambda(x+\hat{\mu}) U_\mu^\dagger(x) \Lambda^\dagger(x)\,. \tag{1.30b}$$

Now the objective is to construct the gluonic part of the lattice Lagrangian, using link variables and respecting the gauge-invariance under transformations Eq. (1.30). The simplest object fulfilling the latter is the trace of the path-ordered product of links around a unit square in the (μ,ν)-plane (cf. Fig. 1.3), the so-called *elementary plaquette*

$$U_{\mu\nu}(x) = U_\mu(x) U_\nu(x+\hat{\mu}) U_\mu^\dagger(x+\hat{\nu}) U_\nu^\dagger(x)\,. \tag{1.31}$$

Defining $G_{\mu\nu}(x) = G_{\mu\nu}\left(x + \frac{1}{2}(\hat{\mu}+\hat{\nu})\right)$ on the lattice, close inspection shows that

$$U_{\mu\nu}(x) = \mathbb{1} + i a^2 g G_{\mu\nu} - \frac{a^4 g^2}{2} G_{\mu\nu} G_{\mu\nu} + \mathcal{O}(a^6)\,. \tag{1.32}$$

Considering that there are six positively oriented plaquettes on a four-dimensional lattice and taking traces on both sides, we obtain

$$\frac{6}{g^2} \sum_x \sum_{\mu<\nu} \left(1 - \frac{1}{3} \mathrm{Re}\,\mathrm{Tr}\, U_{\mu\nu}\right) = \frac{a^4}{2} \sum_x \sum_{\mu,\nu} \mathrm{Tr}\left(G_{\mu\nu} G_{\mu\nu}\right) + \mathcal{O}\left(a^2\right) \tag{1.33a}$$

$$\overset{a\to 0}{=} \frac{1}{4} \int d^4x\, G^a_{\mu\nu} G^{a\,\mu\nu} = \int d^4x\, \mathcal{L}_G^E = S_G^E\,. \tag{1.33b}$$

Here, we made use of $a^4 \sum_x \to \int d^4x$. We see that we recover the continuum gauge action S_G^E in the Euclidian in the limit $a \to 0$. Therefore, we define (after re-expressing

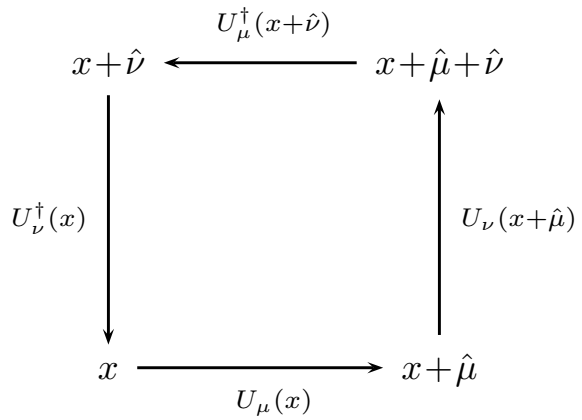

Fig. 1.3: The elementary plaquette.

$\mathrm{Re}\,\mathrm{Tr}\,U_{\mu\nu}$) the Wilson lattice gauge action to be [12]

$$S_{G,L} = \beta \sum_x \sum_{\mu<\nu} \left[1 - \frac{1}{6} \mathrm{Tr} \left(U_{\mu\nu} + U_{\mu\nu}^\dagger \right) \right] , \tag{1.34}$$

where the *lattice coupling* β is defined as

$$\beta = \frac{6}{g^2} . \tag{1.35}$$

From Eq. (1.33a) we know that Eq. (1.34) only holds up to $\mathcal{O}(a^2)$. With the freedom to choose any other discretized action as long as it has the right continuum limit [(1.33b)], so-called *improved actions* [17] are possible. The Lüscher-Weisz action [18–20], for instance, includes more extended loop operators which cancel the lattice artifacts up to $\mathcal{O}\left(a^4\right)$. However, the implementation of improved actions is more involved than that of Eq. (1.34), which is why we confine ourselves to the Wilson formulation.

Ultimately, the goal of this work is to carry out numerical simulations at very high perturbative orders. Unleashing the fermionic degrees of freedom would increase the computational costs to an intolerable degree. As a consequence, we will concentrate on pure gauge theory or *gluodynamics*[6], and only mention in passing some aspects of fermion fields on the lattice. Replacing the derivative in Eq. (1.16) by a symmetrized difference

$$\Delta_\mu^{\mathrm{sym}} \psi(x) = \frac{1}{2a} \left[\psi(x+\hat{\mu}) - \psi(x-\hat{\mu}) \right] , \tag{1.36}$$

[6] Another (slightly more general) characterization for suppressed fermionic vacuum polarisation effects is the *quenched approximation*, in which only valence quarks are allowed.

the naive lattice action reads [12]

$$S_{F,L} = \sum_{f=1}^{n_f} \sum_x \Big\{ \bar{\psi}_f(x) m_f \psi_f(x) + \frac{1}{2a} \bar{\psi}_f(x) \gamma^\mu \Big[U_\mu(x) \psi_f(x+\hat{\mu}) - U_\mu^\dagger(x-\hat{\mu}) \psi_f(x-\hat{\mu}) \Big] \Big\} \,. \tag{1.37}$$

It conforms with Eq. (1.30) and the gauge transformation law for the fermion fields,

$$\psi(x) \to \Lambda(x)\psi(x)\,, \qquad \bar{\psi}(x) \to \bar{\psi}(x)\Lambda^\dagger(x)\,. \tag{1.38}$$

A new problem arises on the lattice, as the naive action Eq. (1.37) suffers from *fermion doublers*, unphysical poles in the lattice fermion propagator due to the discrete nature of the Fourier transform on a hypercube. Doublers can be removed, but only at the price of losing either (exact) lattice chiral symmetry or ultra-locality [21]. Nowadays, many different alternatives to the orginal Wilson formulation are in use, each of them with physical or algorithmic (dis)advantages of their own. For a comparison of the methods, we refer to standard textbooks such as [22] or review articles (e.g., [23]).

2 From Stochastic Quantization to NSPT

With the lattice formalism established in Sec. 1.4, one could in principle start calculating arbitrary QCD observables R by numerically evaluating the path integral Eq. (1.21). It turns out, though, that the configuration space is much too large to evaluate the complete partition function Z, even in pure gauge theory. *Monte Carlo methods* show a way out of this dilemma as they approximate the exact solution by considering only configurations with large statistical weight (*importance sampling*), which is ensured by generating configurations with a probability distribution $\propto \exp(-S_{G,L})$. Different algorithms such as *Metropolis*, *Heat bath* or *Hybrid Monte Carlo* can be used to create configuration sequences (or *Markov chains*) of length N_{MC} in a series of updating steps. Ensemble averages are then approximated as the arithmetic mean over the Markov chain, along with a statistical error that diminishes as $1/\sqrt{N_{\mathrm{MC}}}$ such that

$$\langle R \rangle = \overline{R} \equiv \lim_{N_{\mathrm{MC}} \to \infty} \frac{1}{N_{\mathrm{MC}}} \sum_{i=0}^{N_{\mathrm{MC}}} R(i)\,. \tag{2.1}$$

In the following, we introduce a peculiar evolution in configuration space based on the Langevin equation and discuss the necessary steps to eventually conduct perturbative simulations of lattice gauge theory.

2.1 Stochastic Quantization

The Langevin equation in its original form describes the brownian motion of a particle in a suspension [24],

$$m\ddot{x} = -\gamma\dot{x} + \eta\,. \tag{2.2}$$

The overall force $m\ddot{x}$ on the particle comprises the (deterministic) drift term $-\gamma\dot{x}$ due to friction, and a random force η. This *stochastic noise* embodies the particle's inumerable collisions with the liquid's molecules.

Stochastic Quantization (SQ) [25] generalizes Eq. (2.2) in order to calculate expectation values in field theories and therefore presents an alternative to the path integral method, for example. SQ has a broad application range, reaching as far as the quark-gluon plasma [26], quantum gravity [27] or the sign problem [28]. For reviews of SQ, see [29–31].

To begin with, it shall suffice to consider a scalar field $\phi(x)$ depending on spacetime x

and governed by an action $S[\phi]$. The Langevin equation then reads

$$\frac{\partial\phi(x,t)}{\partial t} = -\frac{\partial S[\phi]}{\partial\phi(x,t)} + \eta(x,t)\,, \tag{2.3}$$

where t is the so-called *stochastic time* that we now let the field ϕ additionally depend on, i.e., $\phi(x) \to \phi(x,t)$.

The noise η yields

$$\begin{aligned} \langle\eta(x,t)\rangle_\eta &= 0\,, \\ \langle\eta(x,t)\eta(x',t')\rangle_\eta &= 2\delta(x-x')\delta(t-t')\,. \end{aligned} \tag{2.4}$$

This definition is equivalent to a Gaussian noise, with probability distribution and generic average over noise itself given by

$$P\left[\eta\right] \propto e^{-\frac{1}{4}\eta^2}\,, \tag{2.5}$$

$$\langle\ldots\rangle_\eta = \int \mathcal{D}\left[\eta\right] P\left[\eta\right] \ldots\ . \tag{2.6}$$

Given an arbitrary observable $R(\phi)$, it can be shown[7] that the average over stochastic time

$$\overline{R(\phi)} = \lim_{T\to\infty}\frac{1}{T}\int_0^T dt\,\langle R\left[\phi(t)\right]\rangle_\eta\,, \tag{2.7}$$

reproduces the expectation value obtained by other approaches, e.g., the path integral method [Eq. (1.21)],

$$\overline{R(\phi)} = \langle R\rangle = \frac{1}{Z}\int \mathcal{D}\left[\phi\right]\ R\left[\phi\right]\ e^{-S[\phi]}\,. \tag{2.8}$$

Hence, SQ provides an independent route to measure physical observables in field theories.

2.2 Numerical Stochastic Perturbation Theory

If we want to exploit the stochastic approach for QCD, the analogue of Eq. (2.3) for the gauge fields must account for their group-theoretical properties. Hence, the Langevin equation is modified to

$$\frac{\partial}{\partial t}U_\mu(n,t) = -it^a\left[\nabla^a_{n,\mu}S_{G,L}[U] + \eta^a_\mu(n,t)\right]U_\mu(n,t)\,, \tag{2.9}$$

where $S_{G,L}[U]$ is the gauge action [Eq. (1.34)] and the generators t^a we already introduced in Eq. (1.3). Again, the sum convention is implied for repeated color indices. In Eq. (2.9), the *Lie derivative* of a function f dependent on U is defined as [34]

$$f\left(e^{it^a\omega^a}U\right) = f(U) + \omega^a\nabla^a f(U) + O(\omega^2)\,, \tag{2.10}$$

[7] For a vigorous proof in perturbation theory see [32], less stringent arguments are given in [33].

with ω being an infinitesimal angle.

A numerical integration of Eq. (2.9) becomes possible by introducing a finite step size ϵ in stochastic time t,

$$t = m\epsilon\,, \tag{2.11}$$

where m is understood as labeling updates and appears in bracketed superscripts in what follows. With the index i numbering all $U_\mu(n)$, the integration of Eq. (2.9) at finite ϵ reads

$$U_i^{(m+1)} = e^{-it^a F_i^a}\, U_i^{(m)}\,. \tag{2.12}$$

The simplest realization of the *force term* F is of Euler type,

$$F_i^a = \left(\epsilon \nabla_i^a S_{G,L}[U] + \sqrt{\epsilon}\eta_i^a\right)\,. \tag{2.13}$$

The explicit expression for F can be found in [35]. Thus, the numerical strategy is to repeatedly update the system using Eq. (2.12). The Euler scheme suffers from a discretization error of $\mathcal{O}(\epsilon)$. Since Eq. (2.8) only holds for continuous stochastice time t, several simulations at different ϵ values and a subsequent extrapolation are necessary to recover the $\epsilon \to 0$ continuum limit. As is obvious from Eq. (2.11), smaller ϵ values are more costly in terms of computer time because they require a larger number of updates m to complete a Langevin trajectory of sufficient length. Besides, rounding errors set a lower bound for ϵ. Yet very large ϵ values are just as problematic: their probability distribution may digress too far from the equilibrium distribution at $\epsilon = 0$, as follows from the Fokker-Plank equation associated with Eq. (2.9).[8] As far as simulation costs are concerned, it is thus beneficial to use higher-order integrators which improve over the $\mathcal{O}(\epsilon)$ precision of the Euler scheme. In our work, we employ a particular integrator which is $\mathcal{O}(\epsilon^2)$ exact and consists of two steps [36],

$$U_i' = \exp\left[it^a\left(\tfrac{-3+2\sqrt{2}}{2}\epsilon\nabla_i^a S[U] - \tfrac{2-\sqrt{2}}{2}\sqrt{\epsilon}\eta_i^a\right)\right]\, U_i^{(m)}\,, \tag{2.14a}$$

$$U_i^{(m+1)} = \exp\left[-it^a\left(\epsilon\nabla_i^a S[U'] + \tfrac{(5-3\sqrt{2})C_A}{12}\epsilon^2\nabla_i^a S[U'] + \sqrt{\epsilon}\eta_i^a\right)\right] U_i^{(m)}\,. \tag{2.14b}$$

This algorithm is one variant of a general set of second-order Runge-Kutta schemes with a purely quadratic dependence on ϵ,

$$y = \text{const.} + b\epsilon^2\,. \tag{2.15}$$

Therefore, Eq. (2.14) presents an alternative to the algorithm described in [35]. The latter requires more computations and an additional copy of the gauge fields. Moreover, it turns out that the coefficient b is significantly smaller when using Eq. (2.14) [36]. Hence, it is

[8] This point is discussed at length in [35].

clearly advantageous to employ Eq. (2.14) instead of the variant given in [35].

The discussion has been restricted to the non-perturbative case so far. In Appendix A.2 we outline how SQ can be used in conjunction with perturbation theory to treat a simple toy model. Perturbative simulations of lattice gauge theories on a computer using the Langevin equation were first proposed in [11] and go under the name of *Numerical Stochastic Perturbation Theory* (NSPT). Calculations up to M loops become possible by a formal weak-coupling expansion of the gluon fields.

Taylor expanding series exponentiation and logarithm, one can conveniently switch between Lie algebra and group, $A \leftrightarrow U$ (cf. Appendix A.3). Group properties are easier to enforce in the algebra, where the expansion reads[9]

$$A = A^{(1)}\beta^{-\frac{1}{2}} + A^{(2)}\beta^{-1} + \cdots + A^{(2M)}\beta^{-M}\,. \tag{2.16}$$

Here, β is the familiar lattice coupling [Eq. (1.35)], which is inversely proportional to the strong coupling constant, $\beta^{-1} = g^2/6 = (2\pi/3)\alpha$. For each order i, the series entries $A^{(i)}$ belong to the Lie algebra,

$$\left[A^{(i)}\right]^\dagger = A^{(i)}\,, \quad \mathrm{Tr}A^{(i)} = 0\,. \tag{2.17}$$

The evolution due to the stochastic differential equation (2.9) is formulated in the group picture with link fields U. The stochastic noise which is characteristic for the Langevin equation enters only at the lowest order $U^{(1)} \propto g$, from which all higher orders follow through a hierarchical system of ordinary differential equations (cf. Appendix A.2, Eq. (A.20), for an illustration in the toy model case). Concretely, a given order $A^{(i)}$ only depends on preceding orders $1, \ldots, i-1$. Therefore, a truncation at finite M is possible.

The perturbative treatment has severe consequences for the computational costs of the simulation. They will be analyzed in detail in Appendix A.4, for the moment we only want to estimate the scaling with the maximal loop order M. Clearly, execution times are driven by the most expensive operation we carry out, the multiplication of two series expansions X and Y:

$$Z = X \times Y \longrightarrow Z^{(j)} = \sum_{i=0}^{j} X^{(i)} \times Y^{(j-i)}\,, \qquad j = 0, \ldots, 2M\,. \tag{2.18}$$

According to Eq. (2.18), the number of necessary matrix multiplications increases in proportion to M^2. Clearly, this scaling behavior is more favorable than, roughly spoken, the factorial growth [37] of the number of Feynman graphs to be evaluated in diagrammatic methods, both in the continuum and on the lattice (for a review focusing on *diagrammatic lattice perturbation theory* (DLPT), see [38]). From this point of view, NSPT is the ideal tool for our perturbative studies at very high orders M, which will be the topic of Chap. 5.

[9] Note the change of notation. The bracketed superscript now labels series entries, not discrete updates.

2.3 Stochastic gauge fixing

QCD being a gauge theory, attached to each physically distinct configuration there is a whole equivalence class of configurations that coincide up to gauge transformations. In perturbation theory, one wants to expand around one given vacuum state. In order to achieve this, the need arises to fix the gauge.

This issue is of fundamental nature and also SQ must deal with it in gauge theories [39, 40]. As can be seen even in the Abelian case of electrodynamics [29], only the physical transversal modes of the gauge fields feel the drift term in the Langevin equation and thus its damping[10] effect. The spurious longitudinal modes, however, are only subject to the noise term and therefore untamed. As far as gauge-invariant observables are concerned, the consequence is that the fluctuations around the equilibrium value grow steadily in the course of Langevin time. *Stochastic gauge fixing* helps in numerical simulations to keep these fluctuations under control, and is a must for gauge-variant observables [41, 42]. The idea behind this procedure is to pull the gauge fields towards the *Landau gauge*,

$$\partial_\mu A_\mu(x) = 0\,. \tag{2.19}$$

On the lattice, this condition can be approached iteratively [43],

$$U_\mu(x) \to e^{i\Gamma(x)} U_\mu(x) e^{-i\Gamma(x+\hat{\mu})}\,, \tag{2.20a}$$

$$\Gamma(x) = -\gamma \sum_\mu \Delta^f_\mu A_\mu(x) = -\gamma \sum_{\mu,a} t^a \Delta^f_\mu A^a_\mu(x)\,, \qquad 0 < \gamma < 1\,. \tag{2.20b}$$

Here, Δ^f_μ is the forward derivative on the lattice and $\gamma \sim \epsilon$ is a convenient choice [43]. Observe that the gauge fields appear in both group and Lie algebra, underlining the importance of the mappings between the two (cf. Appendix A.3). While there may be applications requiring its repeated use to enforce the Landau gauge within numerical precision, for the gauge-invariant observables in our study it suffices to apply Eq. (2.20) once after each Langevin update.

2.4 Zero modes

NSPT faces a problem which also occurs in DLPT on finite lattice volumes [44]: *zero modes* $A_\mu(p = 0)$ are valid solutions of the equations of motion, leading to increasingly divergent fluctuations (NSPT) or singularities within the sum which constitutes the Feynman graph (DLPT). In DLPT, the typical workaround is to neglect the contribution of zero modes [44]. In the limit of infinite lattice volume, this method is expected to produce the

[10] Reconsider Eq. (2.2) and the physical meaning of the drift term.

correct perturbative expansions. In NSPT, one instead makes use of the Fourier transform

$$A_\mu^{(i)}(p) = \int d^4x\, e^{-ipx} A_\mu^{(i)}(x)\,, \tag{2.21}$$

so after each Langevin update one collects the zero modes by a global lattice sum,

$$A_\mu^{(i)}(p=0) = \int d^4x\, A_\mu^{(i)}(x) \to \sum_x A_\mu^{(i)}(x)\,. \tag{2.22}$$

From each of the $N_{\rm Vol}$ links $A_\mu^{(i)}(x)$ in direction μ one then subtracts the proportional share $A_\mu^{(i)}(p=0)/N_{\rm Vol}$ for all perturbative orders i. Also in this context the mapping between group and Lie algebra (Appendix A.3) is useful. Luckily enough, we will encounter a lattice method in Sec. 4.2 that eliminates zero modes completely and makes the subtraction of zero modes superfluous.

2.5 Optimal thermalization of lattices in NSPT

Although NSPT is ideally suited for high-order perturbative calculations, this does not prevent NSPT simulations in this regime from becoming very expensive in terms of computation time. One possibility to reduce it is to *parallelize* the simulation, i.e., distribute it on many cores. The necessity for parallelization, the specific implementations we use and the resulting performance are discussed in detail in Appendix A.4. Below, we want to discuss two further improvements which are more generally applicable in NSPT.

From now on we use the shorthand notation $\{N_S; M\}$ to designate simulations on symmetric volumes N_S^4 up to M loops. The larger the scope of a NSPT simulation, the more expensive it is to *thermalize* a lattice, i.e., to let the initial configuration evolve using Langevin updates until the point when the configuration reaches equilibrium and one can start taking measurements. Here, we present two ways to reduce the computational cost of thermalization:

- *Seed lattice*: If the aim is to run a *farm* of identical simulations to increase statistics, one thermalizes in a first step only one configuration and in a second step branches out into the farm runs continuing the same thermalized lattice but with different random seeds.

- *Accelerated thermalization*: one thermalizes the perturbative orders one after another and thereby ramps up the maximal order on-the-fly until the intended M loops are reached and thermalized.

In the following, we will explain both methods in more detail.

The first method not only has the advantage of minimizing computing time (our largest farms consisted of 200 serial runs for which only one seed lattice had to be thermalized),

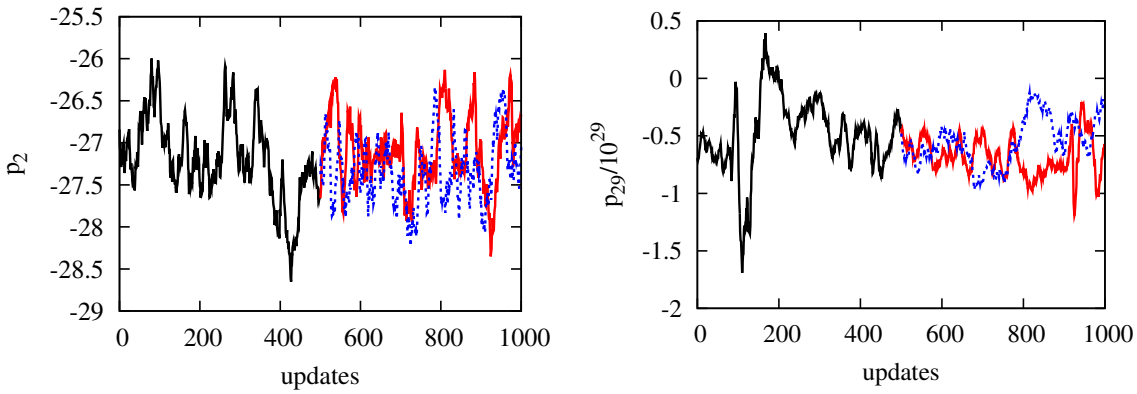

Fig. 2.1: The plaquette coefficients p_2 (left) and p_{29} (right) as defined in Eq. (3.75) from three different $\{10; 30\}$ simulations. Shown are the last 500 of 9770 thermalization steps of one initial lattice (black) and the subsequent 500 updates of two production runs (red, blue) which use the initial lattice to fork off with different random seeds. Note that the overall length of the production runs is $\sim 3 \times 10^4$ to ensure that the vast majority of the data obtained are decorrelated.

especially in conjunction with shared-memory parallelization (cf. Appendix A.4) it significantly reduced time-to-result. What needs to be verified with this procedure is that correlations between production runs due to the joint seed lattice are under control. As Fig. 2.1 illustrates for high and low orders of the perturbative plaquette[11] $\sum_n p_n \alpha^{n+1}$ in $\{10; 30\}$ simulations, this does not pose a severe problem as the Langevin trajectories of production runs decouple fairly soon (and for obvious reasons higher orders lag behind). Furthermore, one can discard a certain number of production measurements or simply make sure that the production runs are much longer than their correlated initial phase. This was always the case for small lattice volumes where the yield of farming is particularly high.

The second method is useful for NSPT simulations in general. We saw that stochastic perturbation theory has a natural hierarchy in the sense that a given perturbative order only depends on the preceeding orders. As long as these are not yet equilibrated, the higher orders will not thermalize either and simply are a waste of computer time. It is even worse, as the disequilibrium of low orders propagates to higher loops which then are thrown further and further from the thermalization point. As a consequence, this conventional thermalization is very expensive.

In place of it, we advocate an accelerated thermalization that significantly cut down our computational costs, as Fig. 2.2 illustrates. The simulation starts with a cold configuration of links of which only the first series entry of $\mathcal{O}(\beta^{-1/2} \sim g)$ is allowed to evolve. As the simulation proceeds, higher perturbative orders in $\beta^{-1/2}$ are successively switched on. In between, we allow $50 - 100$ Langevin updates for each new order to thermalize. This minimizes the risk of lower, unthermalized orders contaminating higher orders. From Fig. 2.2 we conclude that the new procedure reduces the number of updates needed to thermalize each order in $\beta^{-1/2}$ to a constant of about 10.

[11] It will be discussed in detail in Sec. 3.9.2.

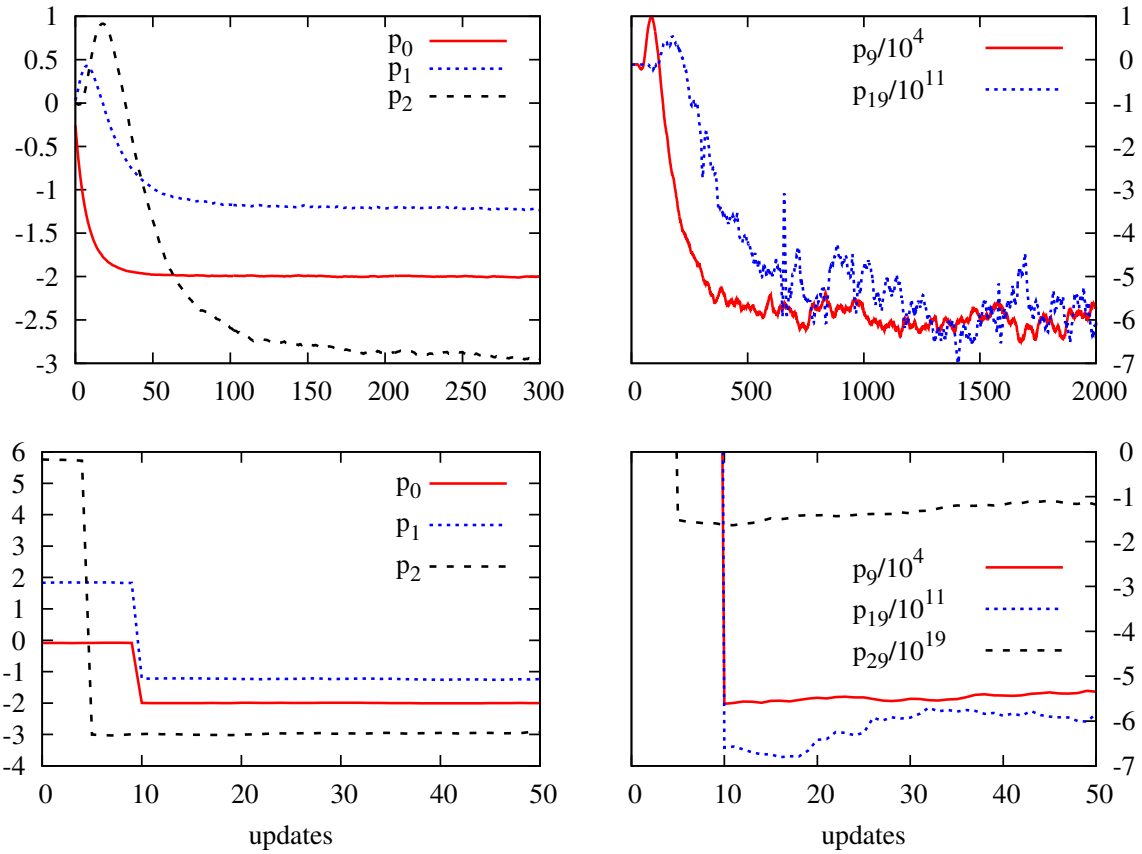

Fig. 2.2: Conventional thermalization with fixed maximal perturbative order (upper row; $\{12;20\}$ simulation) versus accelerated thermalization by ramping up the maximal perturbative order (lower row; $\{10;30\}$ simulation). Shown are the low-order (left column) and high-order (right column) plaquette coefficients, here given in units of inverse lattice β for scaling reasons, approaching thermalization as a function of the completed Langevin updates. The higher the loop order, the more updates are required for conventional thermalization. Instead, accelerated thermalization equilibrates one loop order after another and the required number of updates is only $\mathcal{O}(10)$ per order $\beta^{-1/2}$, independently of the loop order. Therefore, this method greatly reduces the required computer time.

3 Divergent perturbative series

As we will see later, the phenomenon of renormalons in QCD is closely connected to the existence of (at least) two widely separated scales in a calculation: a large and therefore perturbatively accessible scale and a low (energy) scale $\lesssim \Lambda_{\rm QCD}$ such that a perturbative calculation breaks down. A method which tries to bridge this gap is the *Operator Product Expansion* (OPE), and indeed there is a deep connection with renormalon physics, which is why we start out by reviewing the OPE.[12]

3.1 The Operator Product Expansion

Provided that the four-vectors x, y are sufficiently close to each other, the Wilson OPE [47] allows one to rewrite the product of two local fields $A(x)$ and $B(y)$ as

$$A(x)B(y) = \sum_i C_i\,(x-y)\,O_i(x)\,. \tag{3.1}$$

In principle, an infinite number of local operators $O_i(x)$ (with $O_0(x)$ being the unit matrix) may contribute to the expansion, provided that they carry the same quantum numbers as the composite operator. The C_i are complex-valued coefficient functions, the so-called *Wilson coefficients*. Eq. (3.1) is only valid in the weak sense, i.e., once the product $A(x)B(y)$ is evaluated between an initial state $\langle i|$ and a final state $|f\rangle$. Since the operators $O_i(x)$ have corresponding (mass) dimensions d_i,[13] Eq. (3.1) can be regarded as an expansion, ordered in terms of increasing d_i. In the following, we will do the counting in terms of dimensions right away.

Initially, the derivation of Eq. (3.1) focused on Feynman graphs and therefore on its application to perturbation theory [48–50]. Only later it was noted [51–53] that the approach could be extended to include also non-perturbative effects: as a result, a generic observable $R = R(q,\Lambda)$ can be factorized into the (perturbatively calculable) Wilson coefficients $C_d(q,\mu)$ as well as the non-perturbative matrix elements $\langle O_d(\mu,\Lambda)\rangle$ of dimension d:

$$R = C_0(q,\mu)\langle O_0(\mu,\Lambda)\rangle + \sum_d C_d(q,\mu)\langle O_d(\mu,\Lambda)\rangle\left(\frac{\Lambda}{q}\right)^d\,. \tag{3.2}$$

Here, μ denotes a *factorization scale* separating the perturbative scale q and the low

[12] Apart from the original works, this section draws upon several reviews [37, 45, 46].
[13] We repeat that we work in natural units where we can measure everything in powers of mass.

scale Λ, such that $\Lambda \ll \mu \ll q$. The fact that non-perturbative effects take the form of power corrections can be motivated [51] by the *instanton*[14] *density* [54, 55] which scales as $\exp\left[-\text{const}/\alpha\left(q\right)\right]$. The one-loop running of the coupling leads to $\alpha\left(q\right) \sim 1/\ln\left(q^2/\Lambda^2\right)$ [cf. Eq. (A.13)] and therefore gives rise to corrections in powers of $1/q$. The matching of OPE results via dispersion relations to sums over hadronic states leads to what are known as the *QCD sum rules*. This phenomenological approach gained early recognition [45] and today constitutes a vast field of research (e.g., see the reviews [46, 56]) with many different variants [57].

It should be emphasized that in Eq. (3.2) the matrix elements are being evaluated with respect to the *full* QCD vacuum $|\Omega\rangle$, i.e. $\langle O\rangle \equiv \langle\Omega|O|\Omega\rangle$. If Eq. (3.2) were evaluated with respect to the *perturbative* QCD vacuum $|0\rangle$, only the dimension $d = 0$ unit operator would survive, $\langle\Omega|O_0|\Omega\rangle = 1$.[15] Hence, the corresponding Wilson coefficient $C_0 \equiv \sum_n c_n\alpha^n$ would constitute a purely perturbative result for the observable R. But since the full QCD vacuum $|\Omega\rangle$ is not "empty" at all (on the contrary, it is a dense state of strongly coupled quark and gluon fields), there are operators O_d with $d > 0$ whose expectation values with respect to $|\Omega\rangle$ do not vanish, the so-called *vacuum condensates*.[16]

A thorough explanation why condensate is a pertinent name for a basic ingredient of the vacuum can be found in [60]: for instance, one may rewrite the quark propagator as a Taylor expansion given by the free quark propagator plus higher-order corrections, ordered by the number of times the quark scatters from external fields. Within the higher-order corrections, the intricate vacuum structure of QCD permits the replacement of, e.g., one gluon line by two external lines in the following manner: the first gluon vanishes into the vacuum while at the same point in spacetime (or at least at such small distances that the Taylor expansion is still applicable) a gluon spontaneously arises from the vacuum. In that sense, the two external fields "condense" in one point and one attributes them phenomenologically to a certain gluon condensate which we will study later on in great detail.

The structure of the vacuum puts restrictions on the form the operators composing the condensates can take. For instance, they have to be Lorentz scalars and may not carry color charge. It turns out that one cannot construct $d = 1, 2$ operators meeting these criteria. Hence, the condensates with lowest dimension are given by the $d = 3$ quark condensate and the $d = 4$ gluon condensate:

$$O_3 = \overline{\psi}\psi\,, \qquad O_4 = G^a_{\mu\nu}G^{a\mu\nu}\,. \tag{3.3}$$

[14] We have to forestall Sec. 3.6 where instantons are explained.

[15] All other operators vanish in the perturbative vacuum due to normal ordering.

[16] For completeness, we mention recent works [58, 59] claiming that this picture is incorrect and that the vacuum is quite empty indeed. They argue (not only in analogy to but also based on confinement in QCD) that the condensates themselves are contained within hadrons. Consequently, the condensates would represent hadronic properties rather than part of a rich vacuum structure. As not even [58, 59] dispute the phenomenological success of QCD sum rules, we set aside the question where exactly the QCD condensates do originate from.

Beyond that follow the $d = 5$ quark-gluon condensate and the $d = 6$ four-quark and three-gluon condensates. Their operators read

$$\begin{aligned} O_5 &= \overline{\psi}\sigma_{\mu\nu}\frac{\lambda^a}{2}G^{a\mu\nu}\psi\,, \\ O_6^{\psi} &= \left(\overline{\psi}\Gamma_r\psi\right)\left(\overline{\psi}\Gamma_s\psi\right)\,, \\ O_6^{G} &= f_{abc}G^a_{\mu\nu}G^{b\nu}_{\sigma}G^{c\sigma\mu}\,. \end{aligned} \tag{3.4}$$

Here, the $\Gamma_{r,s}$ are placeholders for suitably chosen combinations of color and Lorentz matrices. Due to their power-like suppression, OPE typically do not exceed $d = 6$ and a truncation is possible.

As is evident from Eq. (3.2), both Wilson coefficients and condensates depend on the factorization scale μ, which serves as a renormalization point once these quantities are renormalized. However, if condensates are an adequate description of the physical QCD vacuum, they should be universal in two ways: their shape should neither depend on the very correlator one computes (for the process-dependent piece is contained in the Wilson coefficients), nor on the scheme that is used. Therefore, it is more convenient to replace the above condensate definitions by *renormalization-group invariants* (cf. Appendix A.1). In the case of the quark and gluon condensates, they take the form [61, 62]

$$\langle m\overline{\psi}\psi\rangle\,, \qquad \langle\frac{\alpha}{\pi}G^2\rangle\,. \tag{3.5}$$

Here, the gluon condensate relies on a definition of the gauge action that differs from ours [Eq. (1.15)] by a factor of $1/(4\pi)$ [51–53]. Thus, we would really have $\langle\frac{g^2}{\pi}G^2\rangle$, but the above notation has become the standard. As m is the quark mass, the dimension of the quark condensate is increased by one and the non-perturbative contributions to the OPE start at $d = 4$. The expression for the quark condensate is exact and holds to all orders in perturbation theory (up to perturbative quark mass corrections) while the one given for the gluon condensate reduces the general expression [62] to the massless case.[17] Once estimates for the vacuum condensates are obtained, they can be used for any other observable they appear in, which explains the importance of precise determinations of the condensates. The quark condensate Eq. (3.3) for the lightest quarks was calculated [63] long before the OPE because it is linked to spontaneous chiral symmetry breaking in QCD and yields

$$\frac{1}{2}\langle\overline{u}u + \overline{d}d\rangle(\mu) \approx -\frac{f_\pi^2 m_\pi^2}{2\left[m_{u,\overline{\mathrm{MS}}}(\mu) + m_{d,\overline{\mathrm{MS}}}(\mu)\right]} \backsimeq -\left(254 \pm 15\ \mathrm{MeV}\right)^3\,, \tag{3.6}$$

where f_π is the pion decay constant and m_π the pion mass. The condensate and the

[17] For large quark masses $m \gtrsim \Lambda_{\mathrm{QCD}}$, the gluon condensate mixes with the quark condensate and the respective contributions have to be carefully disentangled in order to avoid double counting [51].

running[18] quark masses $m_{u,\overline{\rm MS}}$ and $m_{d,\overline{\rm MS}}$ are evaluated at the scale $\mu = 2$ GeV in the $\overline{\rm MS}$ scheme [60]. Lattice calculations are in very good agreement with Eq. (3.6): for instance, the estimate of [64] yields $-\left[242(04)(^{+19}_{-18})\ \text{MeV}\right]^3$, also at $\mu = 2$ GeV in the $\overline{\rm MS}$ scheme.

In those OPE that only involve light quarks (see the example in Sec. 3.3), the quark condensate $m\langle\overline{\psi}\psi\rangle$ becomes negligible. In that sense, the gluon condensate is the one with the largest scope. Apart from its phenomenological use, it is of utmost importance for it appears in the *vacuum energy density* [52],

$$\varepsilon = -\frac{\beta_0}{32}\langle\frac{\alpha}{\pi}G^2\rangle\,, \tag{3.7}$$

on account of the trace anomaly [61], and given here in the case without quarks. Thus, a phenomenological value for the gluon condensate can help to put QCD vacuum and hadron models to the test whether they reproduce ε. Later on, we will analyze previous determinations of $\langle\frac{\alpha}{\pi}G^2\rangle$ (Secs. 3.9.2 and 3.10) and make our own attempt in Sec. 5.2.

3.2 The Borel plane

Let us consider a generic observable R in pertubation theory,

$$R \sim \sum_{n=0}^{\infty} r_n \alpha^{n+1}\,. \tag{3.8}$$

Here, the observable R shall start at $\mathcal{O}(\alpha)$ (without loss of generality), and the non-standard indexation will serve our purposes. As we will see later, perturbative series in a theory such as QCD are divergent. At best, they are *asymptotic*. This means that it makes no sense summing the series up to $n = \infty$ and for a given value of α one can identify an order n_0 at which

$$r_{n_0}\alpha^{n_0+1} \tag{3.9}$$

is the minimal term in the series. In a rough approximation, the series Eq. (3.8) should be truncated at n_0 with a truncation error of the size of the minimal term. As far as factorially divergent series are concerned, the accuracy can be improved by applying the *Borel summation*. To this end, one introduces the *Borel transform* of R as

$$B[R](t) = \sum_{n=0}^{\infty} r_n \frac{t^n}{n!}\,. \tag{3.10}$$

[18] This will be explained in Sec. 3.7.1.

Then, provided that Eq. (3.10) has only singularities at real non-positive t and does not grow too fast for $t \to +\infty$, one can perform the *Borel integral* on $B[R](t)$,

$$\tilde{R} = \int_0^\infty dt\, e^{-t/\alpha}\, B[R](t)\,, \tag{3.11}$$

where we rely on α being positive. The crucial point is that $\tilde{R} = R$ perturbatively and, if it can be evaluated, $\tilde{R}$ is the Borel sum of the divergent series Eq. (3.8). As we will see, the requirement that Eq. (3.10) has no singularities at positive t is not fulfilled in QCD (and QED). This has to do with the appearance of fixed-sign divergent behavior in the coefficients r_n in these theories. Therefore, the Borel integral Eq. (3.11) is ill-defined and Borel summation seems useless in this context at first sight. However, the Borel method is still useful as a means to identify and to hierarchize the locations of renormalon poles in the (complex) t plane, the so-called *Borel plane*. Apart from this, sign-alternating divergent series (which we will touch briefly in Sec. 3.5) can indeed be cured by Borel summation, whereas the minimal term approximation does not discriminate between this and the fixed-sign divergence.

3.3 The Adler function as an introductive example

The first observable that has been thoroughly studied [65–70] in view of renormalons and their connection to the OPE [Eq. (3.2)] – and which serves as an ideal starting point for the discussion – is the correlation function of two vector currents,

$$\Pi_{\mu\nu} \equiv (q_\mu q_\nu - q^2 g_{\mu\nu})\, \Pi\left(Q^2\right) = -i \int d^4x e^{iqx} \langle 0|T\{j_\mu(x) j_\nu(0)\}|0\rangle\,. \tag{3.12}$$

Here, we have the *current* $j_\mu = \bar{\psi}\gamma_\mu\psi$, a massless quark field ψ (omitting flavor indices for simplicity) and a large spacelike (Euclidian) momentum $Q^2 = -q^2 \geq 0$. The quantity $\Pi_{\mu\nu}$ is also referred to as the *hadronic vacuum-polarization tensor*, which is of genuine phenomenological importance. For instance, it appears in QCD in the total hadronic cross section in e^+e^- annihilation,

$$\frac{\sigma_{e^+e^-\to \text{ hadrons}}}{\sigma_{e^+e^-\to\mu^+\mu^-}} \propto 12\pi\, \mathrm{Im}\,\Pi\left[\frac{q^2}{\mu^2}, \alpha_s\left(\mu\right)\right] + \mathcal{O}\left(\alpha_{\mathrm{em}}\right)\,. \tag{3.13}$$

The proportionality sign accounts for a constant factor containing the quark charges. It also features in electroweak interaction processes like τ lepton or Z^0 boson decays.

Instead of Eq. (3.12), it is more convenient to address the *Adler function* [71],

$$D\left(Q^2\right) = 4\pi^2 Q^2 \frac{d}{dQ^2} \Pi\left(Q^2\right)\,. \tag{3.14}$$

Up to two loops in perturbation theory (recent determinations reach five loops [5, 72]),

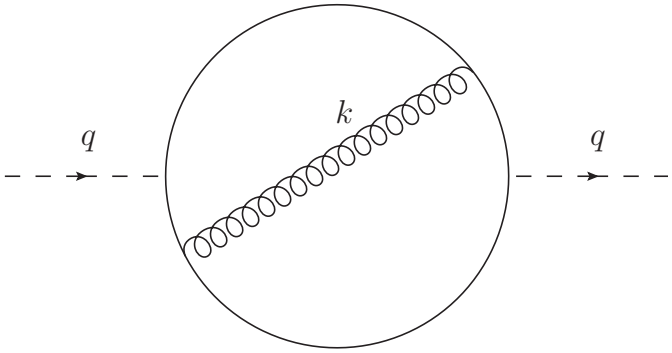

Fig. 3.1: One two-loop Feynman diagram contributing to the vacuum polarization tensor.

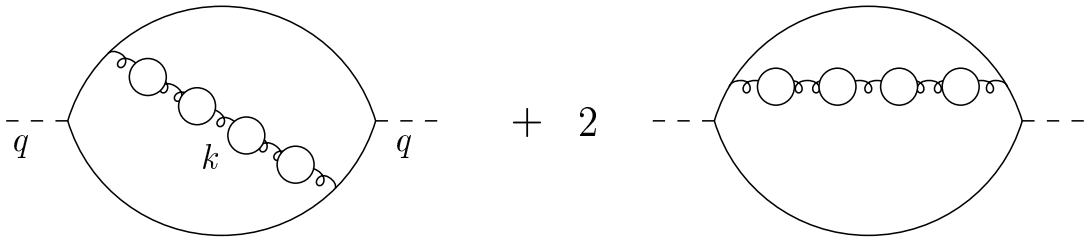

Fig. 3.2: The most simple realizations of bubble chains within the Adler function. Figure taken from [37].

the Adler function yields

$$D_{\mathrm{PT}}\left(Q^2\right) = 1 + \frac{\alpha(Q^2)}{\pi}\,. \tag{3.15}$$

Being the purely perturbative result, this corresponds (apart from the dependence on the factorization scale μ) to the Wilson coefficient function C_0, which was introduced for the OPE Eq. (3.2) and holds at short distances. However, the evaluation of the two-loop diagram requires an integration over the complete range of possible momenta k of a single gluon line exchanged within the fermion loop. It is illustrated in Fig. 3.1 for one example. In the case of a soft (small k) gluon line, the contribution to the k integration is minor but invalid: perturbatively defined propagators should not be employed down to small k where they feel strong non-perturbative effects. Put differently, $k \sim \Lambda$ marks the onset of the non-perturbative regime that prohibits a perturbative approach. So the impossibility to exclude IR regions from perturbative calculations appears at two loops already. It worsens with increasing loop order until it reaches a point where the problem can be seen in the perturbative result itself.

This problem was first noted [73–75] for a certain class of Feynman diagrams with so-called *bubble chains*, multiple insertions of fermion loops into single gluon lines (cf. Fig. 3.2). Effectively, this corresponds to replacing $\alpha\left(Q^2\right)$ by $\alpha\left(k^2\right)$ in the relevant loop integrals. For our purposes it suffices to consider the Adler function Eq. (3.14) in a simplified form [68, 76],

$$\tilde{D}_{\mathrm{PT}}\left(Q^2\right) \sim Q^2 \int dk^2 \frac{k^2 \alpha\left(k^2\right)}{\left(k^2+Q^2\right)^3}\,, \tag{3.16}$$

since it behaves as $D_{\mathrm{PT}}\left(Q^2\right)$ for both $k^2 \ll Q^2$ (IR limit) and $k^2 \gg Q^2$ (UV limit). In these limits, the contributions to the integral Eq. (3.16) yield

$$\begin{aligned} D_{\mathrm{IR}} &= \frac{1}{Q^4}\int_0^{Q^2} dk^2\, k^2 \alpha\left(k^2\right), \\ D_{\mathrm{UV}}^{\mathrm{approx}} &= Q^2 \int_0^{Q^2} \frac{dk^2}{k^4}\alpha\left(k^2\right), \end{aligned} \tag{3.17}$$

where the additional superscript in $D_{\mathrm{UV}}^{\mathrm{approx}}$ will be clarified below. Note that we started out with Q^2 being the only (external) scale, so in order to express Eq. (3.17) in terms of $\alpha\left(k^2\right)$, we can make use of the one-loop running coupling Eq. (A.11) and obtain

$$\begin{aligned} D_{\mathrm{IR}} =& \alpha\left(Q^2\right)\sum_n \left[\frac{\beta_0}{4\pi}\alpha\left(Q^2\right)\right]^n \int_0^{Q^2} \frac{dk^2\, k^2}{Q^4}\left[\ln\left(\frac{Q^2}{k^2}\right)\right]^n \\ =& \frac{\alpha\left(Q^2\right)}{2}\sum_n \left[\frac{\beta_0}{8\pi}\alpha\left(Q^2\right)\right]^n n!\,, \\ D_{\mathrm{UV}}^{\mathrm{approx}} =& \alpha\left(Q^2\right)\sum_n \left[-\frac{\beta_0}{4\pi}\alpha\left(Q^2\right)\right]^n \int_{Q^2}^{\infty} \frac{dk^2\, Q^2}{k^4}\left[\ln\left(\frac{k^2}{Q^2}\right)\right]^n \\ =& \alpha\left(Q^2\right)\sum_n \left[-\frac{\beta_0}{4\pi}\alpha\left(Q^2\right)\right]^n n!\,. \end{aligned} \tag{3.18}$$

The crucial point is that the IR and UV tails of the Adler function, once they are expressed in terms of $\alpha\left(Q^2\right)$, exhibit a factorial growth in their perturbative coefficients. Such a pattern in the perturbative series came to be known as a *renormalon* [75], and according to their origin in the small-k or large-k limit, respectively, one distinguishes between IR and UV renormalons. Apart from the factorial growth they have in common, Eq. (3.18) reveals two important differences between the IR renormalon D_{IR} and the UV renormalon $D_{\mathrm{UV}}^{\mathrm{approx}}$. First, the latter is sign-alternating while the IR renormalon is of fixed sign. Second, the constant that enters the powers of n differs by a factor of 2. At this point, the Borel method of Sec. 3.2 proves useful to classify the two renormalons:

$$\begin{aligned} B\left[D_{\mathrm{IR}}\right](t) &\propto \frac{1}{1-\frac{\beta_0}{8\pi}t} = \frac{1}{1-u/2}\,, \\ B\left[D_{\mathrm{UV}}^{\mathrm{approx}}\right](t) &\propto \frac{1}{1+\frac{\beta_0}{4\pi}t} = \frac{1}{1+u}\,, \end{aligned} \tag{3.19}$$

where $u = \beta_0/\left(4\pi\right)t$. We see that the result Eq. (3.18) is equivalent to *poles* (singularities) in the Borel plane. The IR renormalon pole lies at $u = 2$ while the UV renormalon pole is situated at $u = -1$. We have followed a simple illustration [76] in which the approximation Eq. (3.16) is insensitive to the fact that the UV renormalon at $u = -1$

really is a double pole [70],

$$B\left[D_{\mathrm{UV}}\right](t) \propto \frac{1}{\left(1+\frac{\beta_0}{4\pi}t\right)^2} = \frac{1}{(1+u)^2}\,. \tag{3.20}$$

This corresponds to an additional factor of n in the full D_{UV} [37] compared to the approximation $D_{\mathrm{UV}}^{\mathrm{approx}}$ given in Eq. (3.18).

Before continuing the discussion for the Adler function, several remarks are in place:

1. The above renormalon poles are only the first ones in a series of IR and UV renormalon singularities located at $u > 2$ and $u < 1$, respectively. Therefore, they are referred to as the *leading* IR and UV renormalon poles.

2. This implicit hierarchy is justified as the large-order behavior of the perturbative series is dictated by the singularity closest to the origin of the Borel plane. In that sense, the other UV and IR renormalon poles we neglected above are indeed subleading.

3. It also follows that for $n \to \infty$ the perturbative series will be dictated by the leading UV renormalon since it lies closer to the origin. This can be equally seen from the faster growth of $D_{\mathrm{UV}}^{\mathrm{approx}}$ with increasing n [Eq. (3.18)].

4. However, we shall see in Sec. 3.5 that UV renormalons can be cured in QCD while the unavoidable IR renormalons are conceptually far more interesting.

5. Bubble diagrams are not the only ones giving rise to renormalon divergence as any other diagram related to the running of α will contribute as well.[19]

6. Renormalons are not the only source of divergence in perturbative series. We will briefly touch another source, instanton-anti-instanton pairs, in Sec. 3.6.

Some of the aspects above are visualized in Fig. 3.3.

We will now continue our discussion of the Adler function and solely focus on the IR renormalon. With the one-loop formula Eq. (A.11) for the running coupling we have so far relied upon what is known as the *large-β_0 approximation*. For instance, including β_1 for the IR renormalon of the Adler function amounts to [65, 70]

$$B_{\beta_1}\left[D_{\mathrm{IR}}\right](t) \propto \frac{1}{(1-u/2)^{1+b}}\,, \tag{3.21}$$

where $b = \beta_1/\left(2\beta_0\right)^2$. Hence, the previously simple pole [Eq. (3.19)] does not change its position at $u = 2$ but is joined by a *branch cut* [70], leading to subleading $1/n$ suppressed terms in the asymptotic formula Eq. (3.18). We continue with the large-β_0 limit, but will return to this point in Sec. 3.7.

[19] Nonetheless, most discussions in the literature focus on this type of diagrams for simplicity.

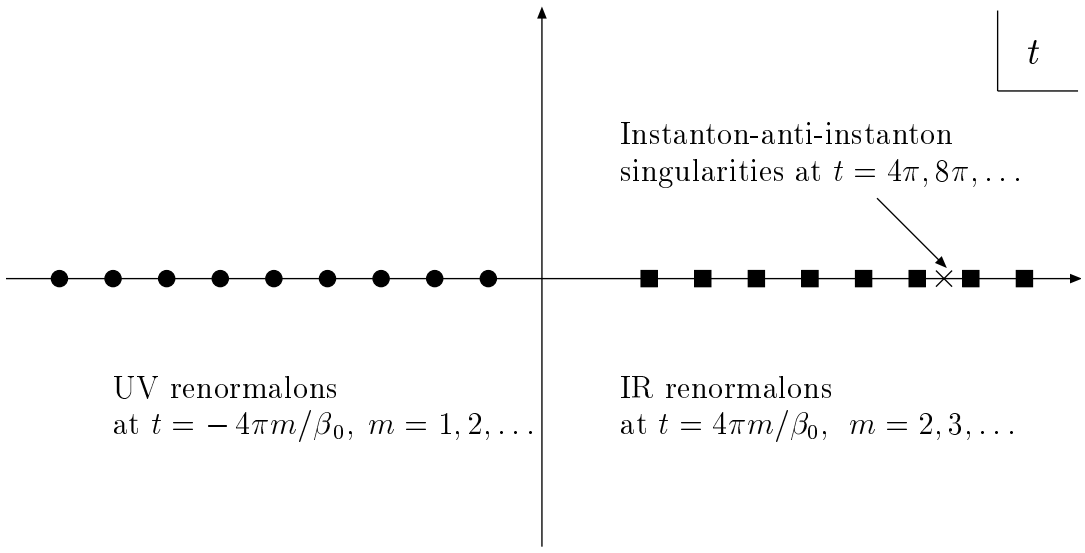

Fig. 3.3: Renormalons as singularities in the Borel plane for the Adler function. Figure taken from [37] and subsequently modified.

In Sec. 3.2 we remarked that an initial approach to asymptotic series is to truncate them at the minimal term, which inevitably leads to a truncation error that is called *renormalon ambiguity* in our context. We shall write our IR renormalon D_{IR} as a perturbative series $\sum p_n \alpha^{n+1}$ and try to locate the order n_0 of the minimal term. Approximately,[20] the minimal term satisfies

$$p_{n_0-1}\alpha^{n_0} \sim p_{n_0}\alpha^{n_0+1}\,. \tag{3.22}$$

Therefore, it yields

$$\frac{p_{n_0+1}}{p_n} \sim \alpha^{-1} \sim \frac{\beta_0}{8\pi} n_0\,, \tag{3.23}$$

where we have made use of the asymptotic formula Eq. (3.18). Hence, the minimal term is located at

$$n_0 \sim \frac{8\pi}{\beta_0 \alpha}\,. \tag{3.24}$$

Inserting this once more in Eq. (3.18), we have

$$p_{n_0}\alpha^{n_0+1} \propto \left(\frac{\beta_0\alpha}{8\pi}\right)^{\frac{8\pi}{\beta_0\alpha}} \Gamma\left(1+\frac{8\pi}{\beta_0\alpha}\right)\,. \tag{3.25}$$

Here, we have employed the Γ-function to account for the factorial of non-integer estimates for n_0. Having convinced ourselves that $(1/x)^x\,\Gamma(1+x) \sim \exp(-x)$ in the conceivable range

[20] A more accurate approach would be to rewrite Eq. (3.18) using Sterling's formula. It would facilitate the differentiation with respect to n as well as the incorporation of β_1 and higher.

of $x = n_0$, we arrive at

$$p_{n_0}\alpha^{n_0+1} \propto \exp\left(-\frac{8\pi}{\beta_0\alpha}\right). \tag{3.26}$$

Λ_{QCD} as defined in Eq. (A.13) marks the low scale in the one-loop running [Eq. (A.11)] such that

$$\alpha\left(Q^2\right) \sim \frac{1}{\frac{\beta_0}{4\pi}\ln\left(Q^2/\Lambda^2_{\mathrm{QCD}}\right)}, \tag{3.27}$$

and therefore

$$p_{n_0}\alpha^{n_0+1} \propto \exp\left[-2\ln\left(Q^2/\Lambda^2_{\mathrm{QCD}}\right)\right] = \left(\frac{\Lambda^2_{\mathrm{QCD}}}{Q^2}\right)^2 = \left(\frac{\Lambda_{\mathrm{QCD}}}{Q}\right)^4. \tag{3.28}$$

As we identify the minimal term with the ambiguity in the perturbative series, we see that the IR renormalon D_{IR} gives rise to an uncertainty inherent in the Adler function $D\left(Q^2\right)$ which scales as

$$\delta D_{\mathrm{IR}}\left(Q^2\right) \propto \left(\frac{\Lambda_{\mathrm{QCD}}}{Q}\right)^4. \tag{3.29}$$

Another possibility [37, 77] to see this would have been to consider the imaginary part of the corresponding Borel integral [Eq. (3.11)]. The advantage of this approach is that it permits a systematic inclusion of β_1 and higher. Still, the ambiguity of the exact order Q^{-4} persists. Let us repeat that the IR renormalon shows a fixed sign divergence which obstructs Borel summation (cf. Sec. 3.2). So the ambiguity that appears in the perturbative calculation [Eq. (3.29)] cannot be circumvented, but yet we saw that the Adler function is of physical relevance and it should be possible to define and calculate it unambiguously.

A way out of this puzzle is being offered by the OPE, which simply tells us that the perturbative result for the Adler function is only one part of the story. To obtain the full result, also non-perturbative effects must be taken into account. Tight constraints [37, 65] can be put on the possible form of the higher dimensional operators:

1. The IR renormalon of the Adler function is due to the low-k regime of the internal gluon line. It is not related to the only external scale Q, which can be factored out. Hence, the low-energy dynamics must be given by a local operator.

2. More precisely, having its origin in the soft gluon line, we also know it must be a gauge field bilinear.

3. With no external hadrons it is clear that we must look for vacuum matrix elements, which, on top of that, are Lorentz scalars like the Adler function.

4. Gauge invariance rules out gauge field combinations such as $A^a_\mu A^{a,\mu}$.

We encountered already the $d = 4$ operator linked to the gluon condensate [Eq. (3.3)]. It is the operator of lowest dimension that meets the above requirements because the contribution of the quark condensate $\langle m\overline{\psi}\psi\rangle$ vanishes in the massless case. The gluon condensate is followed by $d = 6$ operators, such as the three-gluon condensate in Eq. (3.4), which are suppressed by an additional power of $(\Lambda/Q)^2$. So we have

$$D\left(Q^2\right) = D_{\mathrm{PT}}\left(Q,\mu\right) + C_4(Q,\mu)\langle O_4\rangle\left(\frac{\Lambda}{Q}\right)^4 + \mathcal{O}\left[\left(\frac{\Lambda}{Q}\right)^6\right]. \tag{3.30}$$

We see that the first non-vanishing vacuum condensate term is precisely of the order of the minimal term in Eq. (3.29), so up to $\mathcal{O}\left(\Lambda^6/Q^6\right)$ they cancel in the sum. Hence, the Adler function $D\left(Q^2\right)$ is defined unambiguously – it has to be like this since the Adler function is connected to an experimentally accessible observable.

Let us see how well this cancellation works numerically [37, 78]. The gluon condensate's contribution to $\Pi\left(Q^2\right)$ is [51]

$$-\frac{1}{12}\left\{1 + \mathcal{O}\left[\alpha\left(Q^2\right)\right]\right\}\langle\frac{\alpha}{\pi}G^2\rangle\frac{1}{Q^4}\,. \tag{3.31}$$

Higher terms can be found in [79, 80]. The Wilson coefficient C_4 for the specific case of the Adler function can be computed with the help of Eq. (3.14). It amounts to

$$C_4 = \frac{2\pi\alpha}{3}\left\{1 + \mathcal{O}\left[\alpha\left(Q^2\right)\right]\right\}. \tag{3.32}$$

The complete contribution of the gluon condensate then reads

$$\frac{2\pi^2}{3}\langle\frac{\alpha}{\pi}G^2\rangle\,\frac{1}{Q^4} \sim \frac{0.08\,\mathrm{GeV}^4}{Q^4}\,, \tag{3.33}$$

where the original estimate [52] for the gluon condensate is used. By contrast, the ambiguity due to the IR renormalon pole yields

$$\Delta\left(D_{\mathrm{IR}}\right)\left(Q^2\right) \sim \frac{0.06\,\mathrm{GeV}^4}{Q^4}\,. \tag{3.34}$$

As is emphasized in [76], both terms should not cancel exactly, rather the gluon condensate contribution is expected to be numerically larger. The upshot is that the renormalon emerging in the perturbative calculation forebodes the necessity to include non-perturbative effects.

3.4 Generalization to observables of arbitrary dimension

The previous discussion of IR renormalons in the case of the Adler function can be generalized: consider a generic observable R with the first non-vanishing vacuum condensate in

the OPE being of dimension d, then

$$R = C_0(Q,\mu)\langle O_0(\mu,\Lambda)\rangle + C_d(Q,\mu)\langle O_d(\mu,\Lambda)\rangle\left(\frac{\Lambda}{Q}\right)^d + \cdots . \tag{3.35}$$

Likewise, the ambiguity in the perturbative series C_0 will be of $\mathcal{O}\left(\Lambda^d/Q^d\right)$ and cancel with the non-perturbative condensate contribution. As a consequence, the minimal term of the perturbative series will occur at order

$$n_0 \sim d\frac{2\pi}{\beta_0\alpha}\,, \tag{3.36}$$

in one-to-one correspondence with a divergence pattern of the perturbative coefficients

$$p_n \propto \left(d\,\frac{\beta_0}{2\pi}\right)^n\, n!\,. \tag{3.37}$$

One may ask what other observables besides the Adler function can be examined in view of IR renormalons and, in particular, in which of them renormalons can be spotted most easily. From the above discussion we conclude that it is optimal to start the search in $d = 1$ observables: the renormalon divergence should set in most rapidly (lowest n_0) and, correspondingly, lead to the maximal ambiguity of $\mathcal{O}(\Lambda/Q)$ in the perturbative series.

In Sec. 3.7 we will analyze one such $d = 1$ candidate in depth. Before, let us embark on two introductory tours of UV renormalons (Sec. 3.5) and instantons (Sec. 3.6).

3.5 Ultraviolet renormalons

For the Adler function (cf. Sec. 3.3) we saw that two distinct types of renormalons may arise, one of them callled UV renormalon which orginates in the region of large internal momenta. While IR renormalons show a fixed-sign divergence, are non-Borel summable and of physical interest as they can be linked to power-suppressed non-perturbative effects, quite the opposite is true for UV renormalons. This argument is theory-specific and only applies to QCD. In QED, it is the other way round.

The reason is that QED, being an Abelian theory, has a Landau pole in the running coupling in the infrared, i.e., the coupling $\alpha_{\mathrm{QED}}\left(\mu^2\right)$ vanishes for $\mu^2 \to 0$. The non-Abelian structure of QCD leads to the opposite behavior and we discussed in Sec. 1.3 that this manifests itself in opposite signs of the first coefficient β_0 of the β-function (cf. Appendix A.1). Furthermore, we know from Sec. 3.3 that renormalon positions in the Borel plane are determined by β_0 only, and a convenient measure for the position is in terms of $u = t\beta_0/(4\pi)$. From this it is clear already that UV (IR) renormalons, which sit to the left (right) of $u = 0$ in QCD, switch sides in QED.

Studies within QED (e.g., the photon propagator [81, 82]) were important to develop the theoretical tools because the analysis of the sets of diagrams leading to UV renormalon

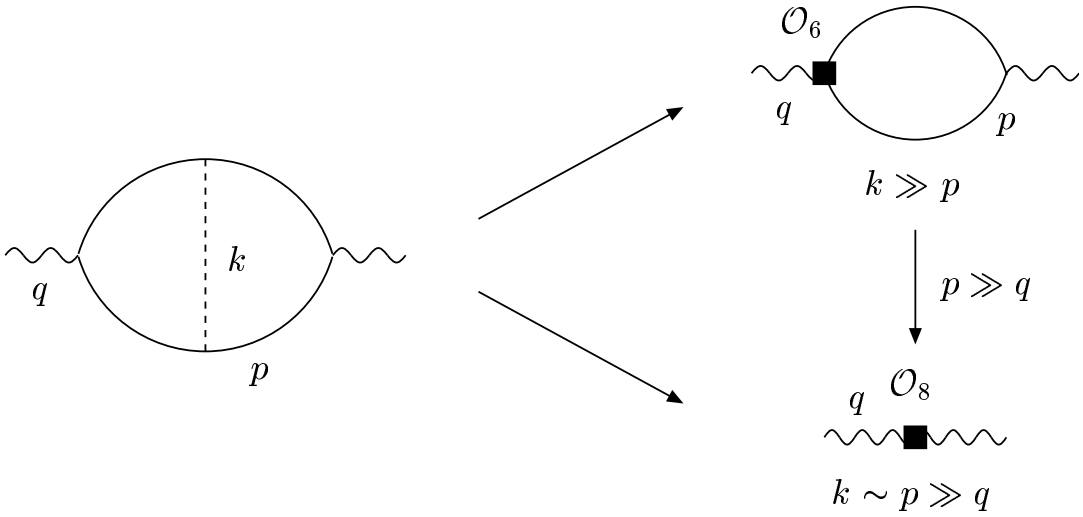

Fig. 3.4: Illustration of the necessary insertions of local operators [right; $\mathcal{O}_6$ corresponds to Eq. (3.38)] at leading order in the $1/n_f$ expansion for the photon vacuum polarization (left). The momentum p flows in the fermion loop. Figure taken from [37].

divergence is much simpler then. A convenient way to tackle the UV problem is given by the *flavor expansion*, where one expands the calculation in $1/n_f$, the number of massless fermion flavors in the theory. In the limit of $n_f \to \infty$ this leads to simplifications for the relevant Feynman diagrams, as those with chains are suppressed by powers in $1/n_f$. In studies of multi-chain diagrams one sees [83], for instance, that higher terms in $1/n_f$ only affect the strength of UV renormalon poles but not their position.

The case of QCD is more complicated, as a glance at $\beta_0 = 11 - \frac{2}{3}n_f$ reveals: its n_f-dependent contribution carries opposite sign. Hence, in a flavor expansion one loses asymptotic freedom at leading order in the running of the coupling already. Indeed, if one restricts the study of bubble chains in QCD to fermion loops only, this leads to UV renormalon poles at positive u, in analogy to QED. By including the next term in $1/n_f$, one recovers the non-Abelian part of β_0. This shifts the renormalon positions as mentioned above. Nonetheless, if applied with caution, the flavor expansion proves useful also in QCD (cf. the review [37]).

We now turn to the question why UV renormalons in QCD are not as severe as their IR counterparts. The answer goes back to a conjecture [84] regarding ϕ^4 theory which is based on renormalization-group arguments. It suggests that the UV renormalons can be remedied by zero-momentum insertions of local gauge-invariant $d = 6$ operators. This argument was put on more solid ground by [83, 85, 86]. As UV renormalons arise for virtual momenta k much larger than the external momenta, $k^2 \gg Q^2$, one can construct an OPE for the photon vacuum-polarization in powers of k^{-2}. What emerges is a complete set of $d = 6$ operators, such as (up to a constant)

$$\partial^\nu F_{\mu\nu} \left(\bar{\psi}\gamma^\mu\psi\right) , \tag{3.38}$$

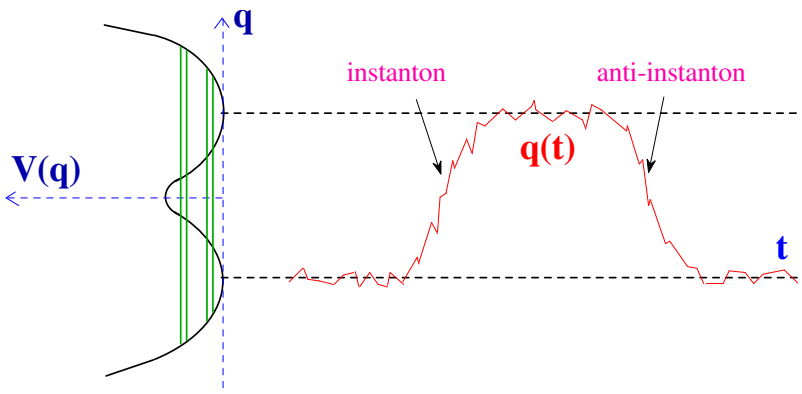

Fig. 3.5: Forward (backward) tunneling as an instanton (anti-instanton) in the one-dimensional double-well potential. Figure taken from [87].

where $F_{\mu\nu}$ is an external abelian (photonic) field strength tensor. These operators can be fine-tuned to serve as local counterterms in an effective four-fermionic Lagrangian that makes the leading UV renormalon divergence disappear (cf. Fig. 3.4). The procedure can be extended to systematically remove all higher UV renormalon poles, and also to account for multiple bubble chains in Feynman diagrams.

We can specify our initial statement that UV renormalon divergences are Borel summable in QCD and therefore do not pose a severe problem: they do give rise to a power-like correction of $\mathcal{O}\left(Q^{-2}\right)$ at leading order [68], but from the above discussion we know now that such corrections are physically irrelevant. This brings us back to Sec. 3.1 and the impossibility to construct $d = 1, 2$ operators for the (physically meaningful) vacuum condensates.

The treatment of UV renormalons is universal since, as soon as one specific process has been analyzed, the UV renormalons of other processes with the same Borel structure can be cured in the same way [83]. The analogy with the (standard) renormalization of UV divergences that arise when evaluating integrals over large loop momenta in fixed-order Feynman diagrams is apparent. Also, the procedure resembles the Symanzik improvement [17] of lattice actions via insertions of higher-dimensional (irrelevant) operators, which we briefly mentioned in Sec. 1.4.

3.6 Instantons

To demonstrate that renormalons are not the only cause of factorial divergence in perturbation theory [88], we take a brief look at so-called *instantons* [54, 89, 90]. The double-well potential in quantum mechanics serves us as an illustration of instantons (Fig. 3.5). In this system they represent the solutions which permit the tunneling of the barrier separating the two minima. It is noteworthy that they cannot be obtained from a purely perturbative expansion around either of the minima.

The analogy with non-Abelian gauge theory arises if one considers, e.g., the QCD vacuum as one such minimum. In fact, there is a whole sea of disjunct QCD vacua. The barriers

between them are of topological nature and can be quantified in terms of different *topological charges*. Transitions are forbidden classically, but have a non-zero amplitude in quantum theory. To be specific, instantons are semi-classical solutions of the equations of motion evaluated in Euclidian spacetime. (Anti-)instantons carry topological charge $(-)1$ and are understood as pseudo-particles because they are localized both in space and time.

Just as in the example above, perturbative calculations in QCD are expansions around an isolated vacuum and are therefore blind to these transitions. However, instanton-anti-instanton pairs with zero topological charge can contribute to the Feynman diagrams [91], giving rise to poles in the Borel plane located at u values which are integer multiples of β_0 [92] in the case of the Adler function. As Fig. 3.3 illustrates, the instanton-related poles lie much farther away from the origin of the Borel plane than the renormalon poles and therefore are subleading as far as the asymptotic behavior of the perturbative coefficients is concerned. Not only the structure of Borel poles, but also the normalization of the perturbative coefficients (*residue* in the Borel language) can be calculated [91], whereas this turns out to be very difficult for renormalons. Another difference with respect to the latter is the signature for instanton-related divergence in Feynman diagrams. Their origin cannot be traced back to a specific momentum region in Feynman diagrams, rather it can be found in the rapid growth of the number of diagrams itself for $n \to \infty$.

We refer to the review [93] for an in-depth discussion of instantons in our context and now return to the discussion of renormalons.

3.7 The pole mass

3.7.1 Quark mass definition and the pole mass renormalon

The notion of a uniquely defined mass for a particle is invalid in quantum field theories. To begin with, there is the *bare mass* m_0 that enters the Lagrangian. The bare mass, just like the other bare parameters (e.g., the bare coupling g_0), is defined at a large scale M which is inaccessible to experiment. For renormalizable theories this means that any appearance of the bare mass is accompanied by the cutoff M in a specific but universal manner, $m_0$itself cannot be measured. At scales $\mu \ll M$ a more meaningful mass definition is needed that actually relates to μ.

In the case of QED, the mass of an isolated electron m_e is such a meaningful definition for scales much lower than m_e itself. Nonetheless, it is already an idealization because in nature an electron sits amidst a cloud of virtual photons, partly screening its charge and mass. These conceptual problems are solved by renormalization, and perturbatively one defines m_e as the pole in the electron propagator. The electron mass is a gauge-invariant quantity and can actually be measured in experiment.

This definition carries over to QCD, in which the quark's *pole mass* m_{pole} refers to the pole in the quark propagator. Although some aspects of the following discussion may equally apply to light quarks, by quark we implicitly mean a heavy quark since we will work

in the *static limit*. The difficulty of QCD is that quarks are confined in hadrons and therefore cannot be observed in isolation. Hence, the quark mass truly is a theoretical quantity, leaving the freedom for various mass definitions taylored for the calculation one faces.[21] The pole mass, for instance, is IR-finite and scheme-, scale- and gauge-independent [95–98]. Also possible are short-distance mass definitions such as the $\overline{\text{MS}}$ *mass* $m_{\overline{\text{MS}}}(\mu)$, which follows from the bare mass m_0 after subtraction of UV-divergent parts in dimensional regularization. Hence, the $\overline{\text{MS}}$ mass is not only IR-finite but also insensitive to long-distance physics. It is a *running mass* that will depend on the normalization point μ. The difference between pole mass and $\overline{\text{MS}}$ mass,

$$\delta m = m_{\text{pole}} - m_{\overline{\text{MS}}}(\mu) \,, \tag{3.39}$$

has been named *residual mass* and we expand

$$m_{\text{pole}} = m_{\overline{\text{MS}}} \left(1 + \sum_{n=0}^{\infty} r_n \alpha^{n+1} \right) . \tag{3.40}$$

The coefficients are known up to r_2 [6, 7, 99], but here we are more interested in the asymptotics of the expansion. In the static limit at one loop the residual mass yields [76]

$$\delta m^{(1)} = -4\pi i C_F \alpha\left(\mu^2\right) \int^{\mu} \frac{d^4k}{(2\pi)^4} \frac{1}{(k_0 + i\epsilon)\,(k^2 + i\epsilon)} \,, \tag{3.41}$$

where the cutoff at μ is owed to the linearly divergent behavior of the integral in the static limit [we will take a closer look at the nature of this divergence around Eq. (3.56)]. The integration over k_0 amounts to a δ-function and we arrive at

$$\delta m^{(1)} = 2\pi C_F \alpha\left(\mu^2\right) \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \frac{1}{\boldsymbol{k}^2} \,. \tag{3.42}$$

This one-loop expression can be generalized to the insertion of bubble chains into the gluon line of the one-loop diagram (cf. Fig. 3.6). The effect of chains of various lengths can be summarized by pulling the running coupling of Eq. (3.42) into the integrand,

$$\delta m = 2\pi C_F \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \frac{\alpha\left(\boldsymbol{k}^2\right)}{\boldsymbol{k}^2} \,. \tag{3.43}$$

To express $\alpha\left(\mu^2\right)$ in terms of the integration variable k, we resort once more to the one-loop formula for the running coupling with the variant given in Eq. (A.12) and obtain

$$\delta m = 2\pi C_F \alpha\left(\mu^2\right) \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \frac{1}{\boldsymbol{k}^2} \sum_{n=0}^{\infty} \left[\frac{\beta_0 \alpha\left(\mu^2\right)}{4\pi} \ln\left(\frac{\mu^2}{\boldsymbol{k}^2}\right) \right]^n \,. \tag{3.44}$$

[21] For a good and yet concise overview of the different mass definitions and how they relate to each other, see [94].

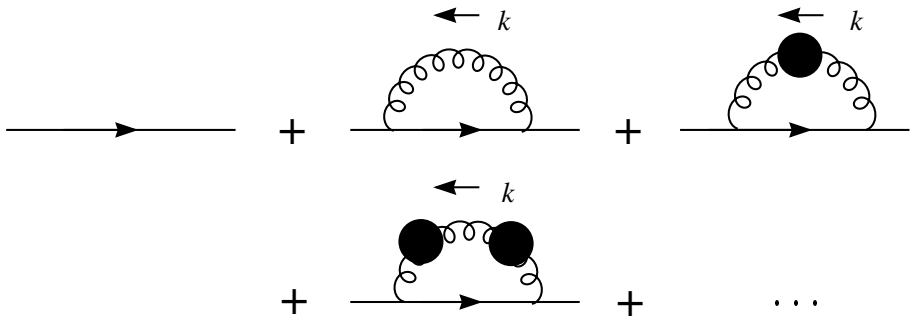

Fig. 3.6: The heavy quark propagator and bubble chains in the one-loop radiative correction. Figure taken from [100] and subsequently modified.

With the help of

$$\int_0^1 dx \left(\ln x^{-2}\right)^n \xrightarrow{n\to\infty} 2^n n! \equiv x_n \,, \tag{3.45}$$

the result reads

$$\delta m = \frac{C_F \alpha\left(\mu^2\right)}{\pi} \mu \sum_{n=0}^{\infty} x_n \left(\frac{\beta_0 \alpha\left(\mu^2\right)}{4\pi}\right)^n = \frac{C_F \alpha\left(\mu^2\right)}{\pi} \mu \sum_{n=0}^{\infty} \left[\frac{\beta_0 \alpha\left(\mu^2\right)}{2\pi}\right]^n n! \,. \tag{3.46}$$

We set the normalization point $\mu = m_{\overline{\text{MS}}}$ and the comparison with Eq. (3.40) reveals

$$m_{\text{pole}} = m_{\overline{\text{MS}}} \left(1 + \sum_{n=0}^{\infty} r_n \alpha^{n+1}\right), \qquad r_n = \left(\frac{\beta_0\left(\mu^2\right)}{2\pi}\right)^n n! \,. \tag{3.47}$$

Now we can apply the machinery we established in Sec. 3.4 for operators of arbitrary dimension d: the growth pattern for the perturbative coefficients r_n matches that of a $d = 1$ operator. At leading order, it follows Eq. (3.37) with d set to 1. The corresponding minimal term occurs at an order n_0 approximately given by Eq. (3.36). By the same token, the ambiguity inherent to δm must be linear in Λ_{QCD}, the Landau pole in the one loop running coupling. Indeed, one finds [76]

$$\Delta\left(\delta m\right) \sim \frac{8}{3\beta_0} \Lambda_{\text{QCD}} \,. \tag{3.48}$$

The range of quoted numbers for the ambiguity is between $\Delta\left(\delta m\right) \approx 50$ MeV ([76], following the above prescription) and $\Delta\left(\delta m\right) \approx (170 - 280)$ MeV ([101], via Borel integral). The discrepancy is not troublesome since it is owed to the differing approaches as well as to the different input values for Λ_{QCD}. It is only important to see that the values lie within the ballpark of Λ_{QCD}, so indeed the IR renormalon Eq. (3.46) we encounter in the pole mass prohibits defining the latter beyond a relative uncertainty $\sim \Lambda_{\text{QCD}}/m_{\overline{\text{MS}}}$. An intuitive explanation of this outcome is that quarks are confined in hadrons with a binding energy $\overline{\Lambda} \sim \mathcal{O}\left(\Lambda_{\text{QCD}}\right)$ [102], which sabotages the determination of the individual quark

masses beyond an uncertainty of this size. Obviously, the relative uncertainty is smallest in the case of heavy quarks, such as the b quark, because $m_b \gg \Lambda_{\mathrm{QCD}}$. We will come back to this point.

3.7.2 The pole mass in HQET

The above renormalon can be approached from a different angle, namely in the context of *Heavy Quark Effective Theory* (HQET) (e.g., [103, 104]; for an exhaustive list, see [100]). The idea of HQET is that the heavy quark is considered static and therefore its degrees of freedom can be eliminated. The heavy quark mass m_Q serves as large scale that cannot be probed by the other (light) degrees of freedom. Effective calculations are then double expansions in α and inverse powers of m_Q. The HQET Lagrangian itself is an expansion and reads

$$\mathcal{L}_{\mathrm{HQET}} = \mathcal{L}_0 + \frac{1}{2m_Q}\mathcal{L}_1 + \cdots + \mathcal{L}_{\mathrm{light}}\,, \tag{3.49}$$

$$\mathcal{L}_0 = \overline{h}_v i v \cdot D h_v\,, \tag{3.50}$$

where h_v is the effective heavy-quark field, v is the heavy quark velocity and $\mathcal{L}_{\mathrm{light}}$ is the effective Lagrangian for the light degrees of freedom. At subleading order in m_Q, the Lagrangian $\mathcal{L}_1$ comprises a number of effective operators such as the "kinetic" term

$$\overline{h}_v \boldsymbol{D}^2 h_v\,. \tag{3.51}$$

Here, $\boldsymbol{D}$ is the covariant derivative restricted to the spatial components. Any physical observable R is then calculated by means of an OPE of the form

$$R = C_0\left(m_Q/\mu\right)\langle O_0\left(\mu\right)\rangle + \frac{1}{m_Q} C_1\left(m_Q/\mu\right)\langle O_1\left(\mu\right)\rangle + \ldots\,, \tag{3.52}$$

with a factorization scale $\mu < m_Q$. To see why HQET is appropriate to address the renormalon in the pole mass, it is instructive to see how the latter is defined explicitly. The full quark propagator in QCD being

$$iS\left(p, m\right) = \frac{i}{\not{p} - m - \Sigma\left(p, m\right)}\,, \tag{3.53}$$

the pole mass m_{pole} is the solution to

$$\not{p} - m - \Sigma\left(p, m\right)\big|_{p^2 = m_{\mathrm{pole}}^2} = 0\,. \tag{3.54}$$

This explains why m_{pole} is also being referred to as the *on-shell mass*. It is the *static quark self-energy* Σ of the heavy quark that order-by-order generates the perturbative behavior of the pole mass (cf. Fig. 3.6). As HQET maintains the low-energy dynamics of QCD by construction, calculating Σ within HQET correctly reproduces [101] the infrared

renormalon encountered in the pole mass for full QCD. More formally, one expects

$$\Sigma = \mu \sum_n c_n \alpha^{n+1} \,, \qquad c_n \xrightarrow{n\to\infty} r_n \,, \tag{3.55}$$

with r_n as in Eq. (3.47). In [101] also the Borel structure of the static energy renormalon is analyzed. IR renormalon-induced double poles are found at positive integer $u \geq 3$, and the absence of IR poles for $u = 1, 2$ (as opposed to the Adler function in Sec. 3.3) can be explained consistently with the absence of the corresponding vacuum condensates. As expected, UV renormalons cause poles at negative u, with the difference that in the HQET formulation appears a $d = 1$ operator, leading to half-integer pole positions. What is entirely new and surprising is the additional appearance of a single UV renormalon pole at positive $u = 1/2$. Borel logic tells us that this UV pole is not Borel-summable. The similarity with the QED case in Sec. 3.5 is misleading. In QED, the UV poles at positive u originate in the opposite sign of the QED β-function compared to QCD. Here, the UV pole at $u = 1/2$ is a peculiarity of HQET. It is caused by the linearly UV-divergent self-energy of the heavy quark. This is understandable if one regards the heavy quark in the static limit as a static source at rest, with a color charge giving rise to a potential of Coulomb-type $\sim \frac{\alpha(1/r)}{r}$. The integration of its Fourier transform over all spatial momenta is of the functional form

$$\int d\boldsymbol{q}^3 \frac{1}{\boldsymbol{q}^2} \sim \int q^2 dq \frac{1}{q^2} \sim \int dq \,. \tag{3.56}$$

We note the linearly UV divergent nature of the integrated color potential, which merges into the static energy.[22] Staying within the picture, one sees how this relates to the IR renormalon in the pole mass: At distances as large as $r \sim \Lambda_{\mathrm{QCD}}^{-1}$, the Coulomb picture for the color field becomes invalid, yet the pole mass definition implies using it for $r \to \infty$.[23] The unjustified use introduces an error of $\mathcal{O}(\Lambda_{\mathrm{QCD}})$.

These findings have strong implications for HQET. As stated above, HQET calculations are expanded in inverse powers of the heavy quark mass m_Q, for which one must choose an adequate prescription first. In [106] it is noted that the physics described by the HQET Lagrangian [Eq. (3.49)] remains the same if one adds to the latter the term

$$-\delta m \overline{h}_v h_v \,, \tag{3.57}$$

containing the residual mass δm we encountered earlier. The heavy quark expansion parameter and the residual mass δm are dependent (choosing a different expansion parameter $m_Q \to m_Q + \Delta m$ readjusts $\delta m \to \delta m - \Delta m$). Therefore, [106] conclude that one is free to set it to zero. However, the conclusion to be drawn from [101] is that the term

[22] This argument is pursued more thoroughly in [105].

[23] The pole mass definition attributes all of the color potential to the heavy quark. In reality, it will be confined: for instance, due to a light quark at finite distance which takes a share of the potential.

in Eq. (3.57) indeed is mandatory in a formulation of HQET which aims to be correct beyond perturbation theory. The appearance of an ambiguous quantity in the Lagrangian seems awkward at first glance, but it is due to the fact that the IR renormalon in the pole mass is seen in real QCD. Any renormalon-free choice $m_Q \neq m_{\text{pole}}$ for the HQET expansion parameter comes at the price of a non-vanishing residual mass reminding us of the renormalon in the full theory.

Before we continue let us summarize and highlight some aspects:

1. When related to the $\overline{\text{MS}}$ mass, the pole mass m_{pole} shows a rapidly divergent perturbative series due to a renormalon at $u = 1/2$ in the Borel plane.

2. The pole mass was the first quantity found to be infrared-finite to all orders in perturbation theory but infrared-unstable beyond.

3. Therefore, m_{pole} is a bad expansion parameter in HQET, although it may seem to be the most natural definition at first sight.

4. An explanation is that the notion of the heavy quark mass being *exactly* on-shell is unphysical.

5. The derivation of the pole mass renormalon relies on the static limit, setting $k_0 = 0$ [cf. Eq. (3.42)]. In other words, the characteristic time scales are very large. For that reason, a regular OPE formulation with local condensates, cf. Eq. (3.2), is inapropriate for m_{pole}. This statement is supported by the fact that a suitable $d = 1$ local HQET operator cannot be found [76].

6. The $\overline{\text{MS}}$ mass is defined unambiguously and hence a better expansion parameter. The most convenient definition is to choose the normalization point $\mu = m_{\overline{\text{MS}}}(\mu)$, such that $m_{\overline{\text{MS}}}(\mu) = m_{\overline{\text{MS}}}\left(m_{\overline{\text{MS}}}\right)$, to avoid large logarithms $\ln\left(\mu^2/m_Q^2\right)$. Not for all observables this may be the most appropriate scale, see the semileptonic decay $b \to u$ discussed in [107].

3.8 Renormalon cancellation in continuum HQET

We pointed out already the renormalon cancellation when it comes to the calculation of physically measurable quantities. For any OPE within HQET, the calculation should not be affected by the ambiguity in the pole mass (or any other prescription used as a mass parameter m_Q). For instance, take the mass of the B meson,

$$m_B = m_{\text{pole}} + \overline{\Lambda}_B + \mathcal{O}\left(\frac{1}{m_b}\right) . \tag{3.58}$$

Here, the light quark mass can safely be neglected and $\overline{\Lambda}_B$ is associated with the meson binding energy. A loose analogy of Eq. (3.58) with the OPE of the Adler function in

Sec. 3.3 arises if m_{pole} is regarded as the large scale Q, and we recall that the binding energy $\sim \Lambda_{\text{QCD}}$, so it can be thought of as taking the role of a low-energy matrix element. The mass m_B of the B meson is measurable, so the right-hand side of Eq. (3.58) must be just as well-defined. That requires the cancellation of the pole mass ambiguity of $\mathcal{O}(\Lambda_{\text{QCD}})$, and indeed this cancellation is due to $\overline{\Lambda}_B$ being ambiguous itself if defined via the pole mass.[24] Noting that the ambiguity in $\overline{\Lambda}_B$ is of the order of its own size, one sees that the binding energy defined in this manner is of little physical importance. By now this comes as no surprise: the very concept of the pole mass negates the confinement of the heavy quark and a light quark within a meson.

While Eq. (3.58) is $\mathcal{O}\left(m_b^{-1}\right)$ precise, one is led to ask whether it is possible to set up an OPE for measurable quantities that surpasses this precision. Consider the charmless B meson decay $B \to X_u l \overline{\nu}_l$. The inclusive final state X_u contains a u quark that is approximately massless. The expression for the total semi-leptonic width $\Gamma(B \to X_u l \overline{\nu}_l)$ can be formulated most naturally within HQET as an OPE in inverse powers of the b quark pole mass $m_{b,\text{pole}}$. The reason is that on the quark level this process is dictated by the b quark decay $b \to u l \overline{\nu}_l$, which in turn is approximated by the free quark decay for which the most natural choice is the pole mass. In the spirit of Eq. (3.2), the OPE reads

$$\Gamma(B \to X_u l \overline{\nu}_l) = \Gamma_{\text{PT}} + \mathcal{O}\left(\frac{\Lambda_{\text{QCD}}^2}{m_{b,\text{pole}}^2}\right), \tag{3.59}$$

where, surprisingly enough, there is no term linear in $\Lambda_{\text{QCD}}/m_{b,\text{pole}}$ in spite of the pole mass ambiguity [109].

Let us see how this comes about: the perturbative part Γ_{PT} reads

$$\Gamma_{\text{PT}} = \frac{G_F^2 |V_{ub}|^2 m_{b,\text{pole}}^5}{192\pi^3}\left[1 + \sum_{n=0}^{\infty} \tilde{r}_n \alpha^{n+1}\left(m_b^2\right)\right]. \tag{3.60}$$

It turns out [76, 101] that the renormalon divergence in the coefficients $\tilde{r}_n$ and the renormalon divergence introduced by $m_{b,\text{pole}}^5$ [re-expressed in terms of the running mass with coefficients r_n as in Eq. (3.47)] cancel up to $\mathcal{O}\left(\Lambda_{\text{QCD}}^2/m_{b,\text{pole}}^2\right)$ in Γ_{PT}. The fast growth of the $\tilde{r}_n$ is even explicitly seen in the large-β_0 approximation [110]. The remaining ambiguity is then compensated by terms containing local $d = 2$ HQET operators such that the OPE Eq. (3.59) is well defined overall.

This example enhances our understanding of the pole mass $m_{Q,\text{pole}}$ of a heavy quark: Although it naturally emerges in measurable quantities of great interest [note the appearance of the CKM matrix element V_{ub} in Eq. (3.60)], its $u = 1/2$ renormalon renders the pole mass useless for high-precision calculations that include non-perturbative effects within an OPE formulation. It would not even make sense to deduce $m_{Q,\text{pole}}$ (provided that all other parameters are known) from a measurement a heavy quark appears in, and

[24] See [76, 101, 106, 108] for a more detailed discussion and possible work-arounds.

use it as an estimate for the "physical" mass of the heavy quark. The poor perturbative behavior makes the estimate loop-specific and causes large theoretical uncertainties due to the perturbative ambiguity. The latter even overshoot the true uncertainties because they must compensate the pole mass renormalon (consider the $\tilde{r}_n$ above). Regardless of these downsides, one encounters quantities such as the one-loop pole mass in the literature (e.g., [111]).

Apart from the two examples given above, there are several other cases in the literature (e.g., [112–114]) with renormalon cancellation in HQET.[25] Recalling the form of a generic OPE in HQET [Eq. (3.52)],

$$R = C_0\left(m_Q/\mu\right)\langle O_0\left(\mu\right)\rangle + \frac{1}{m_Q} C_1\left(m_Q/\mu\right)\langle O_1\left(\mu\right)\rangle + \dots , \tag{3.61}$$

one sees that there is a delicate cancellation of an IR renormalon (contained within the Wilson coefficient C_0) and a UV renormalon (contained within the matrix element $\langle O_1\rangle$). It can be argued [114] that this outcome is obvious if one goes back to the original idea of the Wilson OPE [Eq. (3.1)], which was meant to disentangle short-range from long-range physics in case a problem provides at least two widely separated scales. The OPE within HQET, however, discriminates between perturbative and non-perturbative parts with a "soft" factorization scale μ using dimensional regularization. Instead, one should introduce a "hard" factorization scale $\lambda < m_Q$ to prevent the soft region $k < \lambda$ from intruding perturbation theory as well as forbid a UV contamination of the low energy matrix elements. We will investigate the situation of theories with a hard cutoff in the following.

3.9 Renormalons in theories with a hard regulator

3.9.1 HQET on the lattice

The question of renormalons in theories with a hard regulator arose in view of the promising aspects of implementing HQET on the lattice [115]. The connection with HQET is vital since renormalons as perturbative phenomena should be alien to the non-perturbative lattice world. This is only true, however, for lattice extents La that are large enough to fit a wide separation of scales (say, a small scale Λ and a large scale Q) onto a single lattice, $(La)^{-1} \ll \Lambda \ll Q \ll a^{-1}$. Since this is not feasible in practice, one alternative is to simulate only the soft part $L^{-1} \ll \Lambda \ll a^{-1}$ on the lattice. For the hard part $Q \gg a^{-1}$, one relies on the perturbative Wilson coefficients calculated by matching to the continuum theory.

Once the HQET OPE is put together, the lattice formulation does not lead to the renormalons we are by now familiar with. Instead, it gives rise to corrections which scale like

[25] This is a feature not only of HQET but of effective theories in general, as can be seen from the discussion of the four-Fermi effective theory in [113].

powers of a^{-1} [116], the inverse lattice spacing which serves as a (dimensionful) UV regulator λ. This is not a lattice peculiarity. It occurs in any theory with such a hard regulator λ breaking scale invariance, such as Pauli-Villars regularization [117]. Like renormalons, the power divergences lead to non-perturbative ambiguities,

$$\lambda \exp\left[-\frac{2\pi}{\beta_0 \alpha\left(\lambda\right)}\right] \approx \Lambda_{\mathrm{QCD}}\,, \tag{3.62}$$

that do not arise within purely perturbative calculations.

We will henceforth restrict the discussion to lattice regularization and see how the power divergence manifests itself explicitly. On the lattice, the $d = 5$ HQET kinetic energy operator [cf. Eq. (3.51)] can mix with the $d = 3$ ($d = 4$) operators $\overline{h}_v h_v$ ($\overline{h}_v D_4 h_v$) once the latter are accompanied by quadratic (linear) powers of a^{-1} such that the dimensions match. Let us discuss the consequences for the heavy quark propagator, which in Euclidian continuum QCD satisfies

$$\left(\gamma \cdot D + m_Q\right) S_Q(x,y) = \delta_{x,y}\,. \tag{3.63}$$

It can be expanded in HQET as

$$S_Q(x,y) = \exp\left[-m_Q\left(x_4 - y_4\right)\right]\left[S_Q^0(x,y) + \frac{1}{2m_Q} S_Q^1(x,y) + \mathcal{O}\left(m_Q^{-2}\right)\right]. \tag{3.64}$$

In the static limit and to leading order in m_Q^{-1}, the static propagator is given by [116]

$$S_Q^0(x,y) = \frac{1+\gamma^4}{2}\delta(\boldsymbol{x}-\boldsymbol{y}) L_{\boldsymbol{x}}\left(x_4, y_4\right)\,, \tag{3.65}$$

$L_{\boldsymbol{x}}\left(x_4, y_4\right)$ being the path-ordered exponential connecting the space point $\boldsymbol{x}$ at different times x_4, y_4. On the lattice, the exponential amounts to a product of link matrices, which at leading order in perturbation theory are unit matrices. In evaluating the relevant one-loop self-energy diagrams, one obtains [116, 118, 119]

$$S_Q^0(x,y) = \frac{1+\gamma^4}{2}\delta(\boldsymbol{x}-\boldsymbol{y})\left[1 + \alpha\left(c_0 \frac{t}{a} + w\right) + \mathcal{O}\left(\alpha^2\right)\right]. \tag{3.66}$$

In this context, $\alpha = g_0^2/(4\pi)$ refers to the lattice coupling, $t = x_4 - y_4$ and the quantities c_0, w are given by lattice integral expressions:

$$\begin{aligned} c_0 &= -\pi C_F \int_{-\pi}^{\pi} \frac{d^3k}{2\pi^3}\frac{1}{A}\,, \qquad w = \pi C_F \int_{-\pi}^{\pi} \frac{d^3k}{2\pi^3}\frac{1}{A\sqrt{(1+A)^2-1}}\,, \\ A &= \sum_{i=1}^{3}\left(1-\cos k_i\right). \end{aligned} \tag{3.67}$$

So we see that a linear power divergence of order a^{-1} arises in the leading order static

propagator at one-loop perturbation theory on the lattice. The shape

$$1 + \alpha c_0 \frac{t}{a} \approx \exp\left(\alpha c_0 \frac{t}{a}\right) , \tag{3.68}$$

lends itself to identify $-\alpha c_0 t a^{-1}$ as an additional mass term which is generated on the lattice and can be absorbed into the heavy quark mass renormalization [cf. Eq. (3.65)]. Hence, if one computes lattice correlators for a hadron H containing a heavy quark and decay like $\exp\left(-\mathcal{E}_H t\right)$ for large times t, the energy $\mathcal{E}_H \sim a^{-1}$ bares no physical relevance. Only mass differences (from which the additional mass term cancels) are accessible. In this we note a similarity with the term δm we met in continuum HQET (Sec. 3.7.2) and conclude that we have just computed its one-loop equivalent on the lattice. It generalizes to

$$\delta m\left(a^{-1}\right) = \frac{1}{a} \sum_n c_n \alpha^{n+1} . \tag{3.69}$$

The coefficients c_n and how one can obtain them for $n \gg 0$ play a major role in this work and we postpone their discussion to Sec. 5.1.

Perturbatively subtracting δm on the lattice to render matrix elements finite resurrects renormalons, so [120] introduce a non-perturbative lattice formulation $\delta\overline{m}$ for the residual mass that permits defining a "subtracted" pole mass m^s_{pole} free of both renormalon ambiguities and power divergences. If, however, the final goal is to extract a short-distance mass like $m_{\overline{\text{MS}}}\left(m_{\overline{\text{MS}}}\right)$ [121–125], it is better to stay with the perturbative δm and the pole mass,

$$m_{\text{pole}} = m_H - \mathcal{E}_H\left(a^{-1}\right) + \delta m\left(a^{-1}\right) + \mathcal{O}\left(m_Q^{-1}\right) . \tag{3.70}$$

By matching the full QCD propagator to its HQET expansion, one can eliminate the (unknown) HQET expansion parameter m_Q. For the B meson, it yields:

$$\begin{aligned} m_{b,\overline{\text{MS}}}\left(m_{b,\overline{\text{MS}}}\right) = & \quad m_{b,\text{pole}}\left[1 + \sum_n r_n \alpha^{n+1}\left(m_B\right)\right] \\ = & \quad \left[m_B - \mathcal{E}_B\left(a^{-1}\right) + \delta m\left(a^{-1}\right)\right]\left[1 + \sum_n r_n \alpha^{n+1}\left(m_B\right)\right] + \mathcal{O}\left(m_Q^{-1}\right) \\ = & \quad \left[m_B - \mathcal{E}_B\left(a^{-1}\right) + \frac{1}{a}\sum_n c_n \alpha^{n+1}\left(a^{-1}\right)\right]\left[1 + \sum_n r_n \alpha^{n+1}\left(m_B\right)\right] \\ & + \mathcal{O}\left(m_Q^{-1}\right) . \end{aligned} \tag{3.71}$$

The strong coupling α is taken to be in $\overline{\text{MS}}$. Note that this calculation relies on input from various areas: experiment (the meson mass m_B), lattice simulations (the non-perturbative $\mathcal{E}_B$), lattice perturbation theory (the residual mass δm) and continuum pertubation theory

[the conversion between pole and $\overline{\text{MS}}$ mass, cf. Eq. (3.40)]. What is more, the contributions are combined in such a way that two cancellations take place: the linear divergence in the non-perturbative $\mathcal{E}_B \sim a^{-1}$ perturbatively cancels the one due to $\delta m \sim a^{-1}$, and the renormalon growth in the pole mass expansion coefficients r_n is compensated by the renormalon found in δm. Obviously, both cancellations are of approximative nature. For instance, the latter is exact up to the loop order up to which c_n and r_n are known. It also depends on the scheme conversion $\alpha_{\text{L}}\left(a^{-1}\right) \to \alpha_{\overline{\text{MS}}}(m_B)$.

With this method, one finds [125]

$$m_{b,\overline{\text{MS}}}\left(m_{b,\overline{\text{MS}}}, n_f = 2\right) = (4.26 \pm 0.03 \pm 0.05 \pm 0.07)\ \text{GeV}\,. \tag{3.72}$$

For details on the errors and a discussion on how the result compares to other estimates for $m_{b,\overline{\text{MS}}}$, see [94]. It suffices to say that in the above determination the error on δm is a major source of uncertainty, clearly more important than the error due to the truncation of the HQET expansion, $\Lambda_{\text{QCD}}/m_b \sim 1\,\%$.

3.9.2 The plaquette

In Sec. 1.4 we already introduced the plaquette, whose expectation value on the lattice we shall denote by $\langle P\rangle$. It is linked to the Wilson gauge action $S_{G,L}$ [Eq. (1.34)],

$$S_{G,L} \propto \beta\langle P\rangle\,. \tag{3.73}$$

Classically, one finds in the continuum limit $a \to 0$ [Eq. (1.33)]

$$\frac{1}{a^4}\langle P\rangle = \frac{\pi^2}{36}\langle\frac{\alpha}{\pi}G^2\rangle_{\text{L}} + \mathcal{O}\left(a^2\right) \overset{a\to 0}{=} \frac{\pi^2}{36}\langle\frac{\alpha}{\pi}G^2\rangle\,, \tag{3.74}$$

where we have implicitly defined the lattice version of the gluon condensate we know from Sec. 3.1. Eq. (3.74) holds only classically since $G^{\mu\nu}$ or combinations thereof can mix with lower dimensional operators like the unit operator under renormalization [126]. So, instead of Eq. (3.74), the expectation value $\langle P\rangle$ for the plaquette operator reads

$$\langle P\rangle = \sum_n p_n \alpha^{n+1} + C_G(\alpha)\, a^4\, \langle\frac{\alpha}{\pi}G^2\rangle_{\text{L}} + \mathcal{O}\left(a^6\right)\,. \tag{3.75}$$

Throughout, α denotes the bare coupling on the lattice. For our purposes it suffices that the Wilson coefficient function yields

$$C_G(\alpha) = \frac{\pi^2}{36}\left[1 + \mathcal{O}\left(\alpha\right)\right]\,. \tag{3.76}$$

Since P is a gauge-invariant operator, so are the operators appearing in the OPE Eq. (3.75). This explains the absence of a $d = 2$ operator. Therefore, note the structural similarity with the OPE for the Adler function, Eq. (3.30): in both cases, the first non-vanishing

M	N_S	Reference	M	N_S	Reference
5	8	[11]	12	12	[132]
8	8	[43]	16	8	[132]
10	8,24	[133]	20	4,6,8,12	[134, 135]

Table 3.1: Progress in calculating the coefficients p_n of the perturbative plaquette with NSPT up to M loops on symmetric lattice volumes N_S^4.

higher operator is the $d = 4$ gluon condensate. Here, the inverse lattice spacing takes the formal role of the large scale, $Q = a^{-1}$.

The first three coefficients of the perturbative elementary plaquette in the weak-coupling limit using Wilson gauge action[26] have been known since long, either exactly (p_0 [128]) or with high accuracy from diagrammatic lattice perturbation theory (p_1, p_2 [129–131]):

$$\begin{aligned} p_0 &= \frac{4\pi}{3}\,, \\ p_1 &= 5.35500939669412\,, \\ p_2 &= 27.1983205184481\,. \end{aligned} \tag{3.77}$$

With increasing order, diagrammatic techniques become very laborious. Therefore, the coefficients p_n with $n > 2$ have been obtained only in NSPT and only for finite lattice sizes N_S^4. Again we denote these simulations with $\{N_S; M\}$. In fact, the calculation of the perturbative plaquette [43] was among the first applications of NSPT. Since then, there has been continuous progress in calculating p_n up to higher and higher loops M and yet larger volumes with NSPT (cf. Table 3.1 for an overview).

The unprecedented number of known coefficients has triggered several attempts to study the conjectured plaquette renormalon and the associated gluon condensate. The first serious attempt [43] to answer the renormalon question indeed reported a growth seen in the first eight p_n on a 8^4 lattice that appeared to be close to what was to be expected on theoretical grounds for a $d = 4$ operator (cf. Sec. 3.4). Yet a later analysis [136] of the same $\{8; 8\}$ data to single out the gluon condensate by taking the difference

$$\langle P(\alpha)\rangle - \sum_{n=0}^{n_{\max}} p_n \alpha^{n+1}\,, \tag{3.78}$$

pointed towards a $d = 2$ condensate instead of the expected $d = 4$ (cf. Fig. 3.7). To get a better idea of *finite-size effects* (FSE), some of the previous authors confronted their previous $\{8; 10\}$ calculation with another calculation using $\{24; 10\}$ [133]. They concluded that, up to that order, the FSE lie within 0.5 % of the values for p_n on the 24^4 lattice. Also, they reaffirmed their claim of seeing a factorial growth in the p_n. The disarray grew

[26] For other discretizations and unquenched results, see [127], for instance.

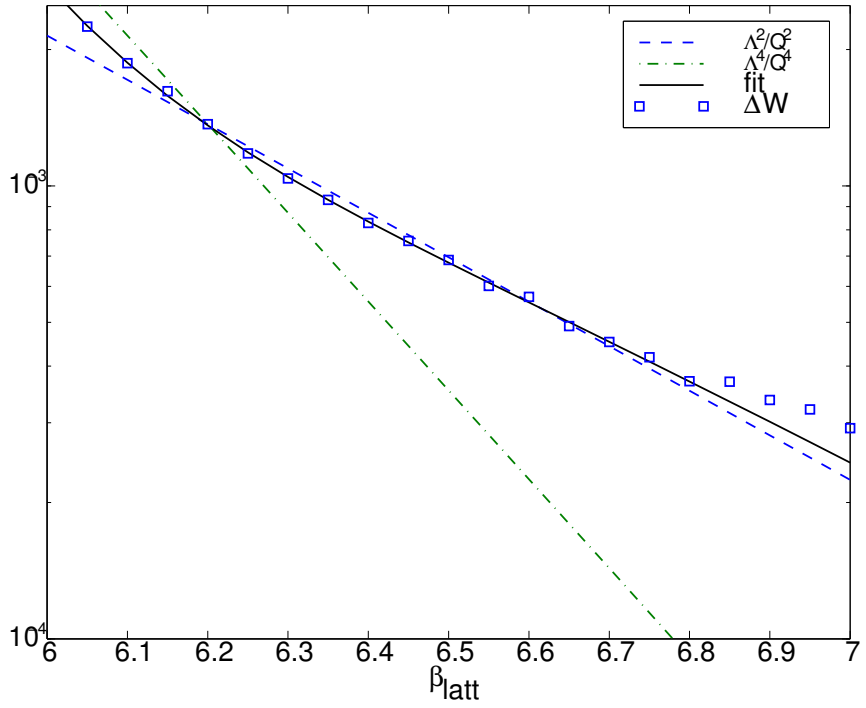

Fig. 3.7: Evidence for a $d = 2$ gluon condensate in the plaquette OPE. Shown is the difference [Eq. (3.78)] between a non-perturbative lattice calculation of the plaquette and the resummed perturbative series at different couplings $\beta = 2\pi/3\alpha$, where a $\{8;8\}$ NSPT calculation was performed. Figure taken from [137].

with [132], now reporting evidence of a $d = 4$ condensate in the plaquette (cf. Fig. 3.8), contradictory to the findings of [136]. In the continuum limit $a \to 0$ they obtain for the gluon condensate [132]

$$\langle \frac{\alpha}{\pi} G^2 \rangle_{(\mathrm{a})} \approx 0.04(1)\ \mathrm{GeV}^4\,. \tag{3.79}$$

Later studies by the same group [134, 135] also dissent from the previous reports [43, 133, 136] in so far as they do not find evidence for a factorial growth in the p_n. Furthermore, they exploit new 20-loop results to concretize the estimate Eq. (3.79). To quote their latest value [135]:

$$\langle \frac{\alpha}{\pi} G^2 \rangle_{(\mathrm{b})} \approx 0.028(3)\ \mathrm{GeV}^4\,. \tag{3.80}$$

Here, an extrapolation of p_n beyond 20 loops relied on a heuristic model in the form of a simple geometric series [138], which disregards the conjectured renormalon dominance at high orders. At any rate, the latest determinations [Eqs. (3.79) and (3.80)] are in agreement with a possible $d = 4$ condensate and, again, they compromise the $d = 2$ condensate seen in [136]. [139] remark that, firstly, this discrepancy might be due to a (too) small maximal loop order in [136]. Secondly, in case the picture of a $d = 4$ condensate in the plaquette persists, the question remains why there is no corresponding renormalon growth in the coefficients p_n despite the one-to-one correspondence of the two concepts in renormalon

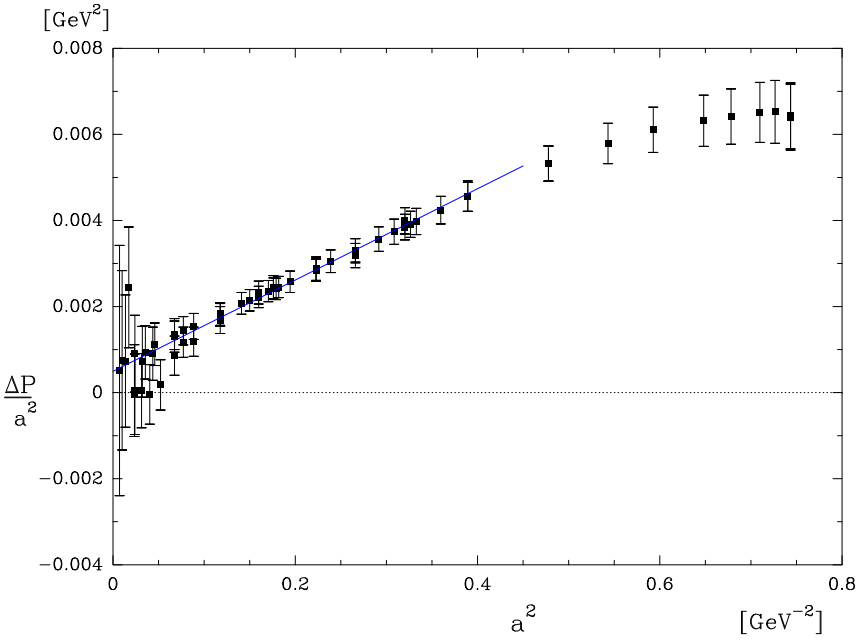

Fig. 3.8: Evidence for a $d = 4$ gluon condensate in the plaquette OPE. Shown is $\langle \frac{\alpha}{\pi} G^2 \rangle_{\rm L}$ from the difference [Eq. (3.78)] between a non-perturbative lattice calculation of the plaquette and the resummed perturbative series as a function of a^2, where a $\{8; 16\}$ NSPT simulation has been used. Figure taken from [132].

theory (cf. Sec. 3.4).

We conclude the discussion by emphasizing that the NSPT studies summarized above have been limited in both maximal loop order M and lattice size N_S without exception,[27] not to mention further (probably subdominant) errors such as the neglect of sea quarks. It is puzzling to see that in all works a stringent extrapolation in N_S has been omitted and yet an extrapolation of the finite volume coefficients p_n for $n \to \infty$ has been carried out in some cases.[28] We will resume this discussion in Sec. 5.2 with our own NSPT study of the plaquette.

3.10 The gluon condensate from phenomenology

As one of two possible condensates [Eq. (3.5)] with lowest dimension, the gluon condensate plays an eminent role in almost any application of the QCD sum rules. It turns out, however, that phenomenological determinations of the gluon condensate are quite difficult.[29] There is a consensus that the first computation [52],

$$\langle \frac{\alpha}{\pi} G^2 \rangle|_{\rm SVZ} = 0.012 \text{ GeV}^4 \,, \tag{3.81}$$

[27] This point is raised in nearly all of the works cited above.
[28] For a discussion of different extrapolations in n for the quenched plaquette see also [140].
[29] The situation is worse than for the quark condensate because the latter can be better isolated from the effects of higher perturbative terms.

from charmonium sum rules has severe drawbacks:

- Experimental input data have improved since then.
- Perturbative corrections were only considered up to NLO in α.
- The previous point is linked to the much too low assumption of $\alpha(m_c) \sim 0.2$, common at the time of [52].
- Higher condensates, in particular O_6^G, were neglected.
- The charmonium sum rules were evaluated at $Q^2 = 0$, thereby increasing the effect of the discarded condensates of higher dimension on the OPE.

Since then, many attempts have been made to improve upon the result Eq. (3.81). The estimates claiming the highest precision come from sum rule analyses of e^+e^- experiments [141],

$$\langle \frac{\alpha}{\pi} G^2 \rangle|_{e^+e^-} = 0.019 \pm 0.002 \text{ GeV}^4 \,. \tag{3.82}$$

Other determinations from e^+e^- data have less predictive power due to strong correlations with other condensates [142, 143]. Another relatively precise computation of the gluon condensate results from the mass splittings of heavy quarkonia [144],

$$\langle \frac{\alpha}{\pi} G^2 \rangle|_{QQ} = 0.024 \pm 0.008 \text{ GeV}^4 \,. \tag{3.83}$$

Eqs. (3.82) and (3.83) are in good agreement and indicate a significant increase compared to the initial estimate Eq. (3.81). However, the different decay channels in the τ decay offer several ways to extract the gluon condensate [145], resulting in a rather low value:

$$\langle \frac{\alpha}{\pi} G^2 \rangle|_{\tau} = 0.006 \pm 0.012 \text{ GeV}^4 \,. \tag{3.84}$$

The poor precision compared to other methods results from an additional α^2 suppression of the gluon condensate in the τ decay. Therefore, its applicability for an accurate determination of $\langle \frac{\alpha}{\pi} G^2 \rangle$ has been criticized [146]. Nonetheless, further evidence of a rather low value of the gluon condensate comes from charmonia [147, 148].

To summarize the status of phenomenological determinations of the gluon condensate, we remark that the different results are quite inconsistent. Namely there are two opposing parties favoring either high or low values of $\langle \frac{\alpha}{\pi} G^2 \rangle$, and the interested reader is referred to both [145] and [60, 146] for a balanced overview. A drawback that all phenomenological determinations have in common is that correlations with other condensates or fit parameters can be large, such as the running charm mass $m_{c,\overline{\text{MS}}}$ in charmonium sum rules [149].

Having seen that the gluon condensate is a quite elusive quantity, we want to close with two further examples underlining why the efforts to determine it precisely are justified.

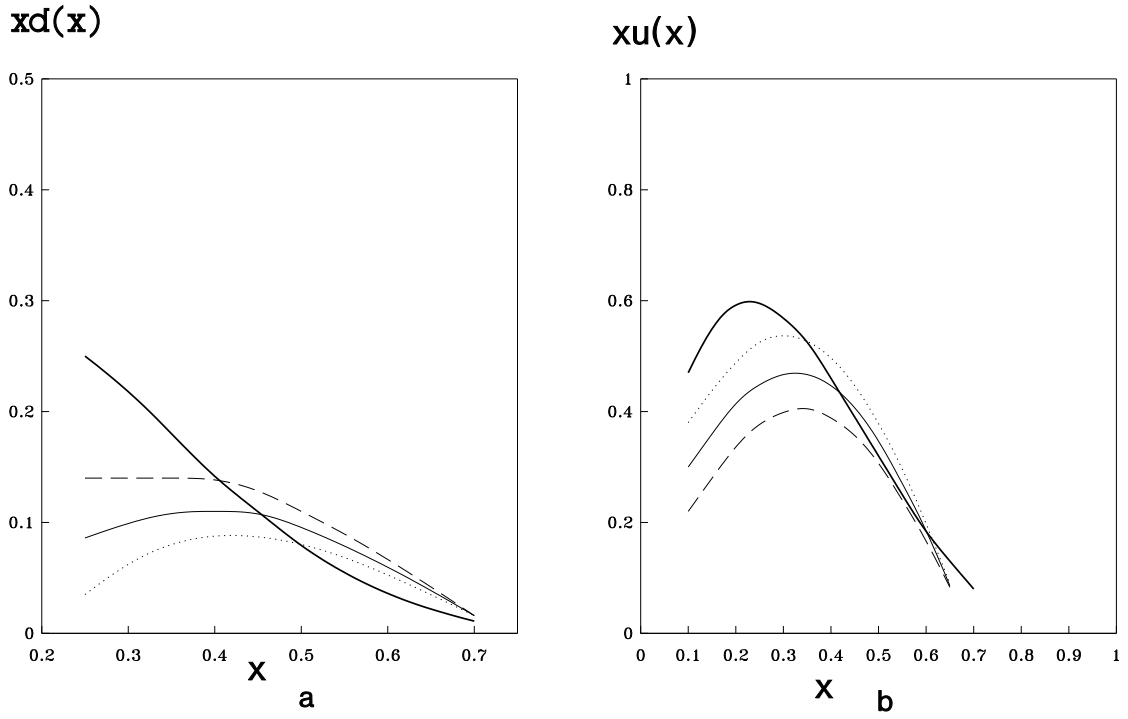

Fig. 3.9: The running d quark (left) and u quark (right) distributions for different values of the gluon condensate $\langle \frac{\alpha}{\pi} G^2 \rangle = 0.012, 0.06, 0$ GeV^4 (dotted, solid and dashed lines, respectively), compared to a leading-order fit ([153], thick solid line) to experimental data from direct photon production. For details, see the original work [154] and the review [145] from which this figure has been taken.

First, with $\langle \frac{\alpha}{\pi} G^2 \rangle$ and two additional input parameters ($m_{c,\overline{\rm MS}}$ and α), one can theoretically predict the electronic decay width $J/\psi \to e^+e^-$ [149]

$$\Gamma^{\rm th}_{J/\psi \to ee} = 4.9 \pm 0.8 \text{ keV} \,. \tag{3.85}$$

This is in fairly good agreement with the measured $\Gamma^{\rm exp}_{J/\psi \to ee} = 5.26 \pm 0.37$ keV. Second, the gluon condensate can serve to set a starting value for quark distributions in nucleons, an information which is precious as the *DGLAP equations* [150–152] can only determine their running after a starting value is set. As an illustration, running d and u quark distributions for different gluon condensate values are shown in Fig. 3.9.

3.11 Toy models and beyond

This section is meant to give a brief overview of models apart from the quantum field theories found in nature that have been studied in view of their large-order perturbative behavior. The reason has been in most cases that the perturbative asymptotics are easier to investigate if technicalities such as renormalization can be avoided.

One such example is the quartic anharmonic oscillator in quantum mechanics with

Hamiltonian

$$H = -\frac{1}{2}\left(\frac{d}{dq}\right)^2 + \frac{1}{2}q^2 + \frac{1}{4}gq^4\,. \tag{3.86}$$

The asymptotic perturbative behavior of its energy levels $E_m(g)$ can be computed analytically and factorial growth is found and confirmed numerically for a generic (polynomial) anharmonic oscillator [10, 155]. It turns out [156, 157] that the factorial growth is not due to renormalons but instantons, which is being signaled (cf. Sec. 3.6) by a corresponding increase of the number I_n of Feynman diagrams that have to be evaluated at given order n. This result generally applies to *super-renormalizable*[30] field theories such as the ϕ_d^4 field theory in $d < 4$ dimensions. In contrast, ϕ_4^4 theory is only renormalizable and the instanton solutions only occur in the massless sector. Nonetheless, they dominate the functional integral in the massive case [158, 159]. Alternating signs of the diagrams ensure instanton-related Borel-summability in the cases of ϕ_d^4 theory discussed above. However, also UV renormalons appear in ϕ_4^4 theory [160], with poles in the Borel plane that lie closer to the origin than the instanton-induced poles. As for any infrared-free theory [75, 161], the UV renormalon poles lie on the positive Borel axis and obstruct Borel-summability. This statement holds also for the inclusion of fermions, as the example of a Yukawa-type interaction shows [162].

In view of asymptotic freedom, the intricate nature of QCD sparks interest in simplified models which still feature this property. For this [163] and other reasons,[31] the O(N) non-linear σ model in $d = 2 - \epsilon$ dimensions with action

$$S = \frac{1}{2}\int d^dx\, \partial_\mu \sigma^a \partial_\mu \sigma^a\,, \tag{3.87}$$

is interesting. Here, the scalar field σ^a, $a = 1, \ldots, N$ fulfills the constraint $\sigma^a\sigma^a = N/g$. The model has been studied in a $1/N$ expansion [9, 66]. The perturbative expansion of Green functions such as

$$G(x, 0) = \langle 0|\sigma(x)\sigma(0)|0\rangle\,, \tag{3.88}$$

exhibits IR renormalons. At leading order in $1/N$, they are located at $4\pi m$, $m \geq 2$ in the Borel plane. Another attempt to mimic the large-order behavior of QCD and yet reduce the complexity is to study QCD with fermionic content, but in two dimensions and the large-N_c limit [164]. Factorial growth is indeed found in the perturbative amplitudes and attributed to instantons.

Investigations of perturbative asymptotics can be carried out as well for theories pointing beyond the Standard Model: perturbation theory for the Bosonic string diverges in a non-Borel-summable fashion [165].

[30] Read: the coupling constant has a positive mass dimension.
[31] See the discussion in [37] for further analogies and limitations.

Theory	**Notation and references**	**High-order behavior**
Anharmonic	$E_m(g) = \sum_{n=0}^{\infty} E_{m,n} g^n$	$E_{m,n} \sim (-3/4)^n n!$
oscillator	[10, 155]	
Anharmonic	$I(g) \sim \int_{-\infty}^{\infty} \mathrm{e}^{(-x^2/2 - g x^4/4)} \mathrm{d}x$	$I_n \sim (-1)^n 4^{-n} (n-1)!$
oscillator, ϕ^4	[93]	
ϕ^4 in 2 and	[166]	Analyticity for complex g,
3 dimensions	[167]	Borel summability proven
Field theories	$Z_n = \sum_{k=0}^{\infty} Z_n^{(k)} (-e^2/(4\pi))^k$	$Z_n^{(k)} \sim k! k^{b_n} A^{-k} c_n [1 + O(1/k)]$
without fermions	[156, 157, 168–171]	
Yukawa theories	[162]	$Z_n \sim n^{-\alpha} A^{-n} \cos(2\pi n/d)$
		$\times \Gamma(n(d-2)/d)$
$d = 2$	[172, 173]	$d_n \sim A^{-n} (\ln n)^n$
O(N) non-linear	[9, 66]	$\sim a^n n!$
σ model		
QED	[92, 174, 175]	$f_n \sim (-1)^n \Gamma[(n+\nu)/2]\, a^{n+\nu}$
	[69] (1st IR renormalon)	$d_n \sim (-\beta_0/(4\pi))^n$
		$\times \Gamma(n + 1 - 2\beta_1/\beta_0^2)$
	[83]	$r_n \sim [\beta_0/(4\pi)]^n\, n!\, n^{2+\beta_1/\beta_0^2+\gamma_2}$
QCD in 2 dim.	[164]	$(g^2 N_c/2)^{2k} (-1)^{k-1} (2k)!$
QCD	[65, 68, 81, 85]	$f_n \sim a^n n^\gamma n!$
	see [69, 83] for discussion	$d_n \sim [-\beta_0/(8\pi)]^n$
		$\times \Gamma(n + 1 - 2\beta_1/\beta_0^2)$
12 Bosonic strings	h is # of handles, [165]	$\sim h!$

Table 3.2: Overview of the asymptotic behavior of perturbative coefficients in different theories. For details and conventions see [176], from which this table has been adopted in part.

The asymptotic perturbative behavior for some of the theories discussed above is summarized in Table 3.2, along with aspects of QED and QCD that were discussed in previous sections.

4 Lattice observables and techniques

4.1 The Polyakov loop

To formulate the Wilson gauge action, we were led to introduce in Sec. 1.4 the elementary plaquette as our first gauge-invariant lattice observable, and discussed it more thoroughly in Sec. 3.9.2. Another type of gauge-invariant observable arises due to the toroidal topology [Eq. (1.26)] we implement on the lattice (see Fig. 4.1). Let us consider asymmetric lattice geometries $N_S^3 \times N_T$ with spatial extent N_S and temporal extent N_T. The *Polyakov loop* is then given by

$$L(N_S, N_T) = \frac{1}{N_S^3} \sum_{\mathbf{n}} \frac{1}{3} \mathrm{Tr} \left[\prod_{n_4=0}^{N_T-1} U_4(n) \right] , \tag{4.1a}$$

$$= \sum_{n=0}^{\infty} l_n(N_S, N_T) \alpha^{n+1} \left(\frac{1}{a} \right) . \tag{4.1b}$$

We are interested in the perturbative expansion of the logarithm of the Polyakov loop,

$$-\frac{\ln\langle L(N_S, N_T)\rangle}{aN_T} = \frac{1}{a} \sum_{n=0}^{\infty} c_n(N_S, N_T) \alpha^{n+1} \left(\frac{1}{a} \right) , \tag{4.2}$$

because it is connected to the self-energy of a static quark on the lattice in the infinite-volume limit [177],

$$\delta m = \lim_{N_S, N_T \to \infty} -\frac{\ln\langle L(N_S, N_T)\rangle}{aN_T} \quad \text{and} \quad c_n = \lim_{N_S, N_T \to \infty} c_n(N_S, N_T) . \tag{4.3}$$

The previous formulae generalize the one-loop considerations of Sec. 3.9.1 and the coefficients c_n coincide with the definition given in Eq. (3.69).

4.2 Twisted boundary conditions

For simplicity, the following discussion shall be restricted to a symmetric lattice with extent La in each direction. Instead of periodic boundary conditions one can impose *twisted boundary conditions* (TBC) [178–182]

$$U_\mu(x + L\hat{\nu}) = \Omega_\nu U_\mu(x) \Omega_\nu^\dagger , \tag{4.4a}$$

$$U_\mu(x - L\hat{\nu}) = \Omega_\nu^\dagger U_\mu(x) \Omega_\nu . \tag{4.4b}$$

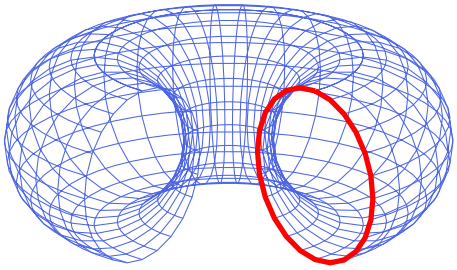

Fig. 4.1: A path of temporal links that spans the complete temporal extent of the lattice, which due to the boundary condition Eq. (1.26) can be visualized by a torus. The path then takes the form of a closed loop.

In other words, when accessing links that sit accross a lattice boundary of a "twisted" direction $\hat{\nu}$, they are multiplied by so-called *twist matrices* Ω_ν that must satisfy

$$\Omega_\mu \Omega_\nu = z\, \Omega_\nu \Omega_\mu, \tag{4.5}$$

$$\Omega_\nu^3 = (-1)^{N-1}\, \mathbb{1}. \tag{4.6}$$

Here, z is a generator of the center Z(3) of SU(3) and therefore $z \in [1, e^{i2\pi/3}, e^{i4\pi/3}]$. From the point of view of a link $U_\mu(x)$, the condition Eq. (4.5) guarantees that the result for the link $U_\mu\left(x + L\hat{\mu} + L\hat{\nu}\right)$ is independent of the order with which two twisted boundaries μ, ν are crossed. The matrices $\Lambda(x)$, which are responsible for gauge transformations [Eq. (1.30)], must obey the same TBC Eq. (4.4).

The lattice gauge action exhibits a discrete symmetry

$$U_\rho\left(x\right) \to \Omega_i U_\rho \Omega_i^\dagger, \tag{4.7}$$

where $\Omega_i = \Omega_\mu, \Omega_\nu, \Omega_\mu\Omega_\nu$, etc. Despite the apparent similarity, Eq. (4.7) is not a gauge rotation [Eq. (1.30)] because $\Lambda\left(x\right) = \Omega_i$ violates Eq. (4.4). Closed Wilson loops (such as the plaquette entering the Wilson gauge action) are invariant under the transformation

$$U_\mu(x) \to z U_\mu(x), \quad \forall x \ni x \cdot \hat{\mu} = \text{const.} \tag{4.8}$$

As a consequence, the SU(3) gauge action is invariant under global transformations Eq. (4.8). This so-called *center symmetry* is exploited when using TBC, whose effects we will discuss shortly.

There are two ways of implementing TBC for the gauge fields. One can keep PBC and multiply the plaquettes in corners of twisted planes with proper factors of z. These plaquettes then contribute differently to the link evolution. The alternative is to impose

Eq. (4.4) with an explicit choice of Ω_ν.[32] We followed the latter route, using

$$\Omega_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \Omega_2 = \begin{pmatrix} e^{-i2\pi/3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{+i2\pi/3} \end{pmatrix},$$
$$\Omega_3 = \Omega_2\Omega_1^2 = \begin{pmatrix} 0 & 0 & e^{-i2\pi/3} \\ 1 & 0 & 0 \\ 0 & e^{+i2\pi/3} & 0 \end{pmatrix}. \tag{4.9}$$

This choice is arbitrary up to unitary transformations: as long as Eq. (4.5) is preserved, the resulting physical amplitudes will not depend on the explicit choice of Ω_μ. TBC in one direction only would be useless as they could be absorbed into a re-definition of the link fields. Therefore, at least two directions must be twisted.

The use of TBC is not restricted to pure gauge theory (in order to incorporate fermions one must resort to TBC plus antiperiodic boundaries, cf. [180, 184]) and the method has proven to be a versatile instrument in many ways: TBC allow the introduction of magnetic monopoles [185] on the lattice [186]. They are regarded as excitations which do not arise dynamically but are forced upon the system, whose response can then be measured, e.g., [178, 187–190]. Furthermore, TBC serve for a modified version [191] of the Eguchi-Kawai model [192], leading to a better approximation of the large-N limit of SU(N) lattice gauge theory in the weak-coupling regime. Lastly, non-perturbative determinations of the running coupling constant on the lattice can be performed using Polyakov loop correlators with TBC [193, 194]. For this work, two particular effects of TBC are beneficial:

1. TBC completely eliminate zero modes (cf. Appendix F of [188]). With PBC, zero modes would require a proper treatment (cf. Sec. 2.4).

2. TBC reduce *finite-size effects* (FSE) as the possible gluon momenta are quantized [181] in integer multiples of

$$k_\nu = \begin{cases} \frac{2\pi}{3La}, & \nu = \text{twisted direction}\,, \\ \\ \frac{2\pi}{La}, & \nu = \text{periodic direction}\,. \end{cases} \tag{4.10}$$

Roughly speaking, the modes in a twisted direction behave as if the corresponding physical lattice extent were $3La$ instead of La, at least to low orders in perturbation theory. To study the effect of TBC quantitatively, we compare lattice results for the two observables which are of interest to us.

In the case of the average perturbative plaquette [Eq. (3.75)], TBC make a huge impact at small lattice sizes, as Table 4.1 illustrates: the one-loop coefficients of both TBC in two (TBCxy) and three (TBCxyz) spatial dimensions on a 4^4 volume are compatible with

[32] See [183] for a recipe how to construct them for arbitrary SU(N).

	4^4 PBC	4^4 TBCxy	4^4 TBCxyz	32^4 PBC	DLPT
$\mathcal{O}(\beta)$	1.9929(1)	1.9998(2)	1.9997(3)	2.0000(2)	2
$\mathcal{O}(\beta^{-2})$	1.2028(5)	1.2185(4)	1.2200(6)	1.2205(6)	1.2207957536
$\mathcal{O}(\beta^{-3})$	2.887(3)	2.954(2)	2.957(2)	2.957(3)	2.9605080512
$\mathcal{O}(\beta^{-4})$	9.05(1)	9.41(1)	9.399(9)		
$\mathcal{O}(\beta^{-5})$	32.49(6)	34.5(1)	34.34(4)		

Table 4.1: Perturbative plaquette coefficients from an expansion in the inverse lattice coupling. TBC data (this work), 4^4 PBC data from [195], 32^4 PBC from [196] and infinite-volume DLPT data with PBC from [131].

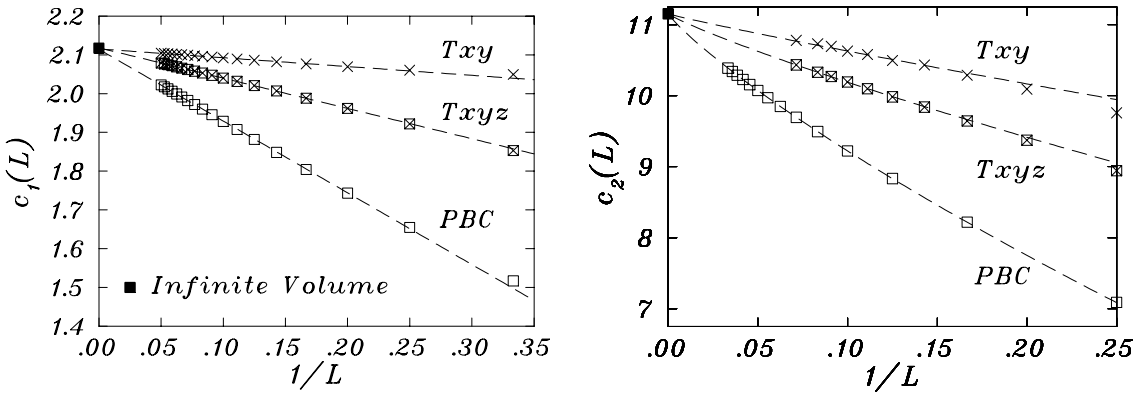

Fig. 4.2: The one and two-loop perturbative coefficient of the static energy extracted from Polyakov loops. Figures taken from [127].

the precise infinite-volume result from DLPT. Also, they agree with 32^4 PBC data at two and three loops. Neither is the case for 4^4 PBC data. Hence, the use of TBC indeed reduces FSE. However, based on the perturbative plaquette it is not possible to discriminate between the options TBCxy and TBCxyz.

The situation is different for our second observable, the Polyakov loop [Eq. (4.1)]. Fig. 4.2 shows the one and two-loop perturbative coefficients of the static energy, extracted from Polyakov loops in Monte Carlo simulations at weak couplings [127] and evaluated for a wide range of symmetric volumes L^4 and different boundary conditions. The latter choice does not have an impact on the extrapolated infinite-volume result, as expected. Furthermore, we see that TBC significantly reduce FSE, leading to a much flatter extrapolation towards infinite volume. If this flatness were taken as the only criterion, Fig. 4.2 would suggest TBCxy to be the boundary condition of choice. However, Figs. 5-7 of [197] illustrate that, in non-perturbative lattice simulations, TBCxy also have a drawback compared to TBCxyz: only the latter truly prevent the simulation from tunnelling between different Z(3) phases (they are due to the center symmetry), while TBCxy merely lead to a reduction compared to PBC. As a consequence, TBCxy simulations fluctuate more and return a noisier signal than their TBCxyz counterpart. We make a similar observation in our NSPT simulations, as Fig. 4.3 illustrates for two simulations on a 16^4 volume at fixed $\epsilon = 0.05$

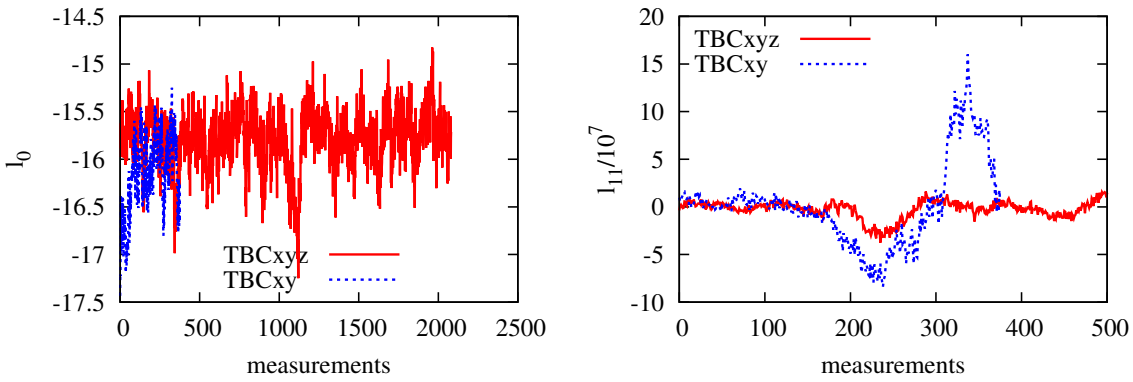

Fig. 4.3: Polyakov loop measurements on a 16^4 lattice. On the left: the one-loop perturbative coefficient l_0 of the Polyakov loop for TBCxyz (blue) and TBCxy (red). On the right: the corresponding l_{11} coefficients, featuring a peak which is symptomatic for the strong fluctuations in TBCxy simulations.

for TBCxy and TBCxyz, respectively. While the measurement trajectories of the one-loop Polyakov loop coefficient l_0 show a similar behavior for TBCxy and TBCxyz, one notes a peak that arises in the twelve-loop l_{11} TBCxy measurement history which is symptomatic for the problems that occur in TBCxy simulations.

4.3 Link smearing

When measuring an observable on the lattice one is typically interested singling out a specific signal (the ground state, for instance) as good as possible. In order to reduce the overlap with unwanted higher excitations, one can resort to *smearing* the link fields (which in turn can be beneficial once quarks are included, see [198–200]). This is to say that a given link is blended or averaged with its surrounding link neighbors in such a way that the resulting "smeared link" can be utilized for an improved determination of the wanted observable. Generally speaking, smearing washes out the higher (and therefore very local) modes, while the low modes in long-distance observables such as correlators come out more sharply. To achieve the desired effect, smearing can be applied once or iteratively. Especially if many smearing steps are necessary, the choice of smearing algorithm matters.

The first available was *APE smearing* [201, 202]. In doing so, a link is blended with the weighted sum $\Sigma_\mu(x)$ of n_{staple} of its six *staples* (a staple is the ordered product of three neighboring links such that the product of the link itself and its staple yields a closed path on the lattice), namely

$$\begin{aligned} U_\mu(x) \to U_\mu^{\text{APE}}(x) &= (1-\alpha_{\text{APE}})U_\mu(x) + \frac{\alpha_{\text{APE}}}{n_{\text{staple}}}\Sigma_\mu(x)\,, \\ \Sigma_\mu(x) &= \sum_{\nu\neq\mu}\Big[U_\nu(x)U_\mu(x+\hat{\nu})U_\nu^\dagger(x+\hat{\mu}) \\ &\qquad + U_\nu^\dagger(x-\hat{\nu})U_\mu(x-\hat{\nu})U_\nu(x-\hat{\nu}+\hat{\mu})\Big]\,, \end{aligned} \tag{4.11}$$

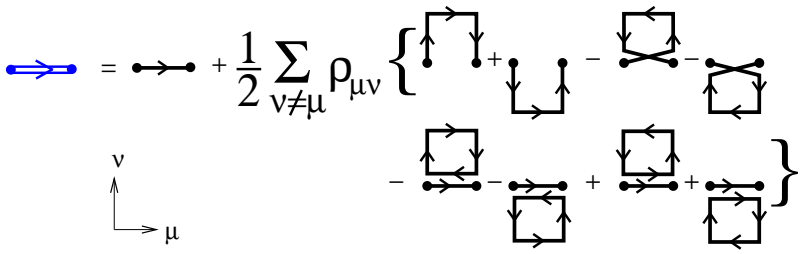

Fig. 4.4: Stout smearing. Figure taken from [203].

where the parameter α_{APE} is the smearing fraction. The apparent drawback of Eq. (4.11) is that the resulting smeared link $U_\mu^{\mathrm{APE}}(x) \notin \mathrm{SU}(3)$ and ought to be projected back onto a group element V, e.g., by iteratively maximizing $\mathrm{Re}\,\mathrm{Tr}\left[U^{\mathrm{APE}}\,V^\dagger\right]$, which may result in increased simulation times. Yet the projection process poses an even bigger problem: APE smearing is not analytic and hence cannot be cast in a form accessible to perturbative expansions. One alternative is *Stout smearing* [203], in which the smeared links are obtained as

$$U_\mu(x) \to U_\mu^{\mathrm{Stout}}(x) = e^{iQ_\mu(x)} U_\mu(x)\,, \tag{4.12}$$

where the multiplicative recipe guarantees that the smeared link stays within SU(3). To this end, $Q_\mu(x)$ must be a member of the algebra by design (cf. Fig. 4.4 for an illustration):

$$\begin{aligned} Q_\mu(x) &= \frac{i}{2}\left[R_\mu^\dagger(x) - R_\mu(x)\right) - \frac{i}{6}\mathrm{Tr}\left(R_\mu^\dagger(x) - R_\mu(x)\right]\,, \\ R_\mu(x) &= C_\mu(x) U_\mu^\dagger(x), \\ C_\mu(x) &= \sum_{\nu\neq\mu} \rho_{\mu\nu}\Big[U_\nu(x) U_\mu(x+\hat{\nu}) U_\nu^\dagger(x+\hat{\mu}) \\ &\qquad + U_\nu^\dagger(x-\hat{\nu}) U_\mu(x-\hat{\nu}) U_\nu(x-\hat{\nu}+\hat{\mu})\Big]\,. \end{aligned} \tag{4.13}$$

Here, the usual sum convention is not implied, and $\rho_{\mu\nu}$ are weights that can be set according to the purpose.

5 Results

5.1 Polyakov loops

The main aspects of this study have been published [204], an extended investigation is ongoing [36]. In essence, our strategy consists of a series of refinements: in Sec. 5.1.1, we specify the formulae of Chap. 3 to the degree we need for our analysis, and devise our setup for the simulations. With the help of Secs. 5.1.2 through 5.1.4, we then sophisticate our original work [204] in Sec. 5.1.5.

5.1.1 Prerequisites

Let us resume the discussion of the pole mass, which so far has mostly been limited to the large-β_0 approximation (cf. Sec. 3.7.1). As in the case of the Adler function, the renormalon ambiguity not only follows from the estimate of the minimal term but also from considering the imaginary part of the Borel integral of $m_{\rm pole}$ [205]. What is appealing about this approach is that it allows to quantify things beyond the large-β_0 limit. The large-order perturbative behavior is dictated by the closest singularity to the origin of its Borel transform, whose precise shape near the closest singularity ($u = 1/2$) can be parametrized as

$$\begin{aligned} B[m_{\rm pole}](t(u)) =& N_m\,\mu\,\frac{1}{(1-2u)^{1+b}}\left(1+s_1(1-2u)+s_2(1-2u)^2+\cdots\right) \\ &+ (\text{analytic term}), \end{aligned} \tag{5.1}$$

where "analytic term" refers to a piece which is expected to be analytic up to the next renormalon ($u = 1$). N_m is a normalization constant and the coefficients b, s_1 [205] and s_2 [37, 206] are exact and read

$$\begin{aligned} b &= \frac{\beta_1}{2\beta_0^2}\,, \\ s_1 &= \frac{1}{4\,b\beta_0^3}\left(\frac{\beta_1^2}{\beta_0}-\beta_2\right), \\ s_2 &= \frac{1}{b(b-1)}\frac{\beta_1^4+4\beta_0^3\beta_1\beta_2-2\beta_0\beta_1^2\beta_2+\beta_0^2(-2\beta_1^3+\beta_2^2)-2\beta_0^4\beta_3}{32\beta_0^8}\,. \end{aligned} \tag{5.2}$$

The perturbative coefficients r_n then behave as

$$r_n \overset{n\to\infty}{=} N_m \left(\frac{\beta_0}{2\pi}\right)^n \frac{\Gamma(n+1+b)}{\Gamma(1+b)} \times \left[1 + \frac{b}{(n+b)} s_1 + \frac{b(b-1)}{(n+b)(n+b-1)} s_2 + \cdots\right] . \tag{5.3}$$

As the energy $E(r) = 2m + V_s(r)$ of a system of two static quarks separated by a distance r is free of renormalon ambiguities, the strength of the renormalon of the singlet static potential V_s must also be determined by N_m such that the contributing renormalons cancel in the energy [105, 207, 208].

Eq. (5.3) replaces the large-β_0 result Eq. (3.47) of Sec. 3.7.1. There, we also saw that the coefficients c_n of the static energy should have the same asymptotics [up to subleading renormalons causing $\mathcal{O}\left(e^{-1/n}\right)$ corrections],

$$c_n \overset{n\to\infty}{=} r_n . \tag{5.4}$$

When computing ratios between two subsequent coefficients, the normalization cancels and the ratios behave as

$$\frac{c_n}{c_{n-1}} \frac{1}{n} = \frac{\beta_0}{2\pi} \left\{1 + \frac{b}{n} - (1 - b s_1) \frac{b s_1}{n^2} - \frac{1}{n^3} \left[b s_1 (b-1) + b^2 s_1^2 (2-b) - 3 s_2 b (b-1)\right] + \mathcal{O}\left(\frac{1}{n^4}\right)\right\} . \tag{5.5}$$

We conduct NSPT simulations (Sec. 2.2) and obtain the coefficients c_n on the lattice by means of the Polyakov loop (Sec. 4.1). In Sec. 4.2, we analyzed the advantages of using twisted boundary conditions in our simulation. As a result, we twist all spatial directions (TBCxyz) and measure the Polyakov loop along the (untwisted) temporal direction.

The Polyakov loop formulae in Sec. 4.1 were given for the fundamental representation, but we can easily generalize them for gauge links $U_\mu^R(n)$ in an arbitrary representation R of SU(3).[33] We also implement the octet or *adjoint* representation with dimension $d_R = 8$ and adjust the normalized trace in Eq. (4.1) accordingly. It gives us access to the equivalent of the self-energy δm for the octet representation. We denote it by δA and the corresponding normalization by N_A. Apart from N_A, the asymptotic expectation for the coefficients [Eq. (5.3)] and the ratios [Eq. (5.5)] is identical to the fundamental representation.

For both fundamental and adjoint sources, we calculate the Polyakov loop with and without Stout smearing (Sec. 4.3) of the temporal links $U_4(n)$. In case of smearing, we set the smearing parameter according to $\rho_{\mu\nu} = 1/6 \;\forall\; \mu$ or $\nu = 4$, the absence of smearing is denoted by $\rho_{\mu\nu} = 0$. Hence, we perform four different measurements altogether and label them with the superscript (R, ρ).[34]

[33] See Eqs. (10) – (17) of [209], for instance.

[34] We will drop the superscript (R, ρ) whenever the strategy is identical irrespective of representation R or link smearing.

	$\mathbf{N_S}$ (N_T)			
$\mathcal{O}(\alpha^3)$	**5**$(5,6,7,8,10)$			
$\mathcal{O}(\alpha^4)$	**4**$(5,6,7,8,10,12,16,20,24)$	**12**$(16,20)$		
$\mathcal{O}(\alpha^{12})$	**6**$(6,8,10,12,16)$	**8**$(12,16)$		
$\mathcal{O}(\alpha^{12})$	**10**$(8,12,16,20)$	**16**$(12,16,20)$		
$\mathcal{O}(\alpha^{20})$	**7**$(7,8)$	**8**$(8,10)$	**9**(12)	**10**(10)
$\mathcal{O}(\alpha^{20})$	**11**(16)	**12**(12)	**14**(14)	

Table 5.1: Overview of Polyakov loop simulations. The first column states to which order in α the coefficients of $c_n^{(R,\rho)}(N_T, N_S)$ and the associated ratios have been computed for each specific lattice volume (second column).

Smearing of temporal links affects the static action $\bar{\psi}D_4\psi$, as the covariant derivative acts on a scalar lattice field $f(n)$ like

$$D_4 f(n) = \frac{1}{2a}[U_4(n)f(n+\hat{4}) - U_4^\dagger(n-\hat{4})f(n-\hat{4})]. \tag{5.6}$$

Since Stout smearing alters $U_4(n)$ in a gauge-covariant fashion, it amounts to a different but equally valid discretization of the static action. Different lattice actions result in different regularization prescriptions that can be distinguished on short ranges (smearing is a local procedure) but do not affect the long-distance physics such as an infrared renormalon. We therefore expect a similar asymptotic behavior [Eqs. (5.3) and (5.5)] for coefficients with smearing and regard it as a universality check.

We have obtained the coefficients $c_n^{(R,\rho)}(N_S, N_T)$ for a large variety of geometries that we list in Table 5.1. The statistical analysis of measurement data was carried out using [210], which allows to process either single runs or jointly evaluate a whole set of independent simulations. Special care was taken that at a given perturbative order every individual history was long enough to guarantee a safe error analysis. If this was not the case, this and all higher orders of the questionable run were discarded.

We will discuss the impact of finite N_S and N_T in detail. First, we analyze how a variation of the stochastic time step affects the resulting c_n.

5.1.2 The influence of the stochastic time step

We employ the second-order scheme (2.14) to integrate the Langevin equation numerically. Therefore, we have to deal with a purely quadratic dependence on ϵ, the size of the stochastic time step.

For all the geometries listed in Table 5.1, we have computed the coefficients at $\epsilon = 0.05$. We will see below that the integrator exhibits a particularly flat extrapolation $\epsilon \to 0$,[35] so

[35] See also the discussion around Eq. (2.15) in Sec. 2.2.

we have effectively simulated at $\epsilon = 0.05 \approx 0$. This constitutes our first set of data for the coefficients $c_n(N_S, N_T)$. We relied upon it for [204] and call it "DATA50".

To estimate the error introduced by this approximation, we have conducted simulations at several different values of ϵ for some of the volumes listed in Table 5.1. We then extrapolated $\epsilon \to 0$ using Eq. (2.15). In Fig. 5.1 we display the outcome for a 16^4 lattice. Shown are unsmeared triplet coefficients $c_n^{(3,0)}(\epsilon)$ for six distinct time steps ϵ, together with the fit curve and the extrapolated result at $\epsilon = 0$. For convenience, each vertical axis is normalized to the $c_n^{(3,0)}(\epsilon = 0.05)$ coefficient. We emphasize the particular flatness of the extrapolated curve in comparison to the size of the statistical uncertainties. This marginal ϵ dependence up to large orders is confirmed by other, smaller volumes for which we simulated at several values of ϵ. There is no apparent volume dependence. It encourages us to assign a generic uncertainty to all those volumes where the extrapolation in ϵ has been omitted. To this end, we obtain the difference between the $\epsilon = 0.05$ result and the extrapolated $\epsilon = 0$ result,

$$d_\epsilon(n) = \left| \frac{c_n\left(\epsilon = 0.05\right) - c_n\left(\epsilon = 0\right)}{c_n\left(\epsilon = 0.05\right)} \right| , \tag{5.7}$$

in the case of the 16^4 lattice. We linearly extrapolate $d_\epsilon(n)$ beyond 12 loops. We then add d_ϵ in quadrature to the relative statistical error of the coefficients on those geometries for which only $\epsilon = 0.05$ results are available. The resulting second set of data is called "DATA0". It differs only slightly from DATA50 for the many lattice volumes where the statistical fluctuations dominate over d_ϵ. In the few cases with very precise results, however, the inclusion of d_ϵ introduces a lower bound on the statistical error and rules out an interference of ϵ effects. All following analyses will be based on DATA0.

5.1.3 The limit $N_T \to \infty$

We isolate the N_S dependence by considering the limit $N_T \to \infty$. Since most of our lattice volumes (cf. Table 5.1) have $N_S \leq N_T$, it is reasonable to take this limit first. In other words, we investigate the self-energy on a finite spatial volume, $\delta m(N_S)$, with coefficients

$$c_n(N_S) = \lim_{N_T \to \infty} c_n(N_S, N_T) \, . \tag{5.8}$$

Regardless of whether one uses PBC or TBC for the lattice boundaries, the static quark propagating in the temporal direction will undergo Coulomb-type interactions with its own mirror charges on replica lattices. The interaction is visualized in Fig. 5.2 for the nearest mirror charge at distance aN_S. The mirror charges give rise to an effective static potential [197],

$$c_n(N_S) = c_n - \frac{f_n(N_S)}{N_S} + \mathcal{O}\left(\frac{1}{N_S^2}\right) . \tag{5.9}$$

We can neglect the N_S^{-2} corrections for sufficiently large N_S.

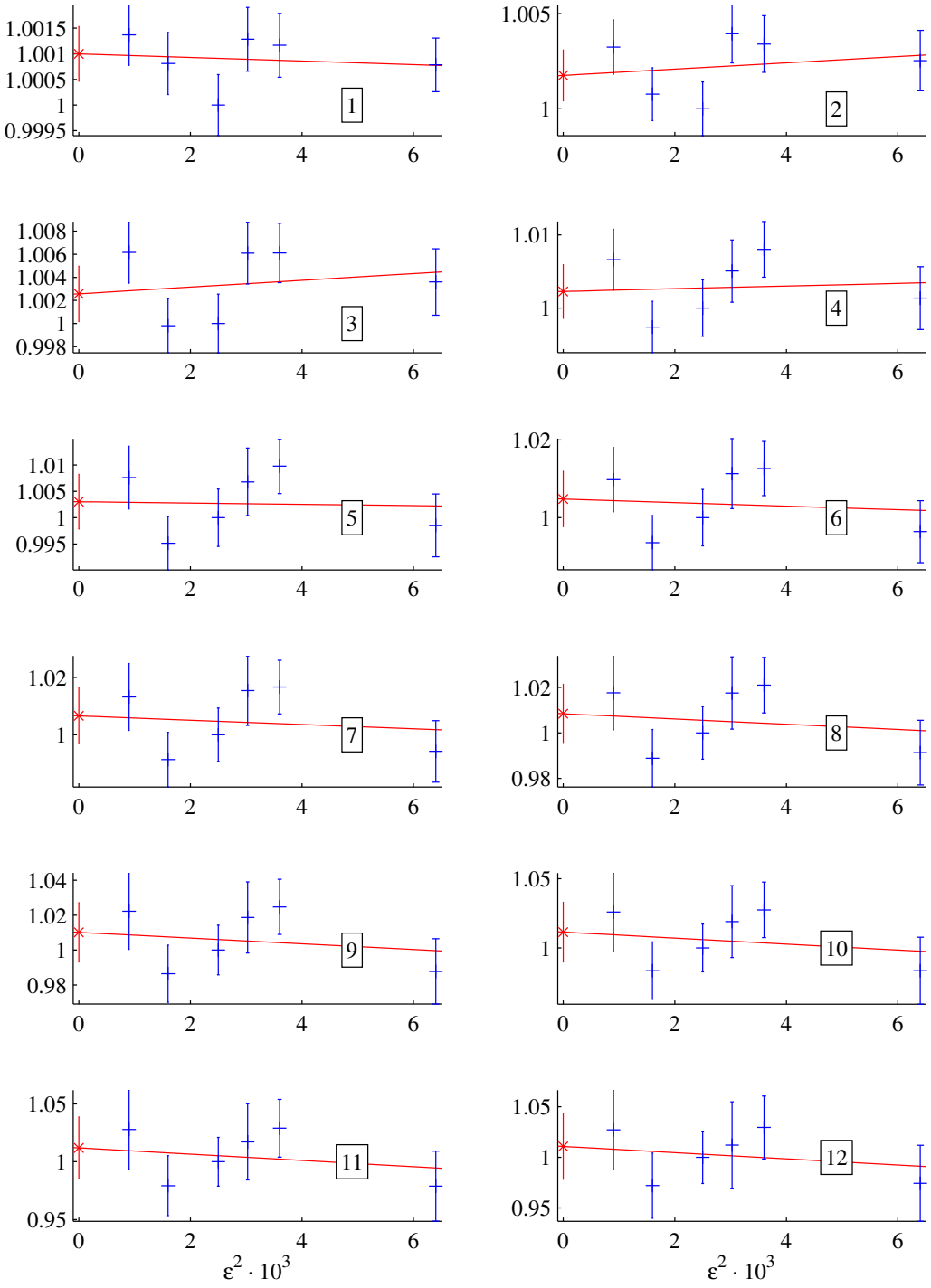

Fig. 5.1: Overview of ϵ effects in simulations on a 16^4 lattice. Numbers in boxes refer to the associated loops. Details are given in the text.

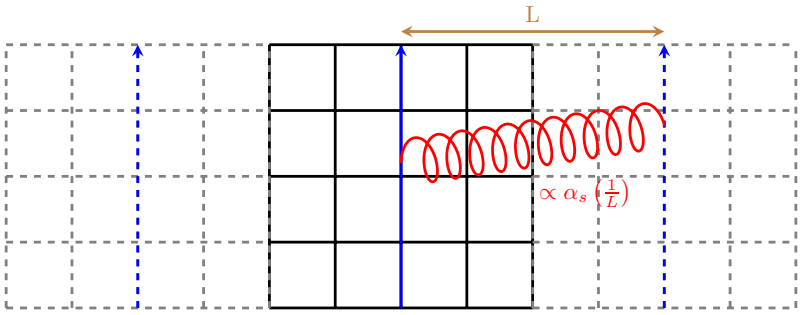

Fig. 5.2: The self-interaction with mirror charges.

Two things are worth noting about this potential: first, it is similar to an interacting part of a potential. In other words, it contains no self-energies, as they are already included in δm. Second, its characteristic scale is $\sim (aN_S)^{-1}$, which leads to

$$\delta m(N_S) = \delta m - \frac{1}{N_S}\sum_{n=0}^{\infty} f_n \alpha^{n+1}\Big[(aN_S)^{-1}\Big] + \mathcal{O}\left(\frac{1}{N_S^2}\right). \tag{5.10}$$

Hence, the term $f_n(N_S)$ in Eq. (5.9) is related to the coefficients f_n by the renormalization group running $\alpha\left[(aN_S)^{-1}\right] \to \alpha\left(a^{-1}\right)$. We parametrize

$$f_n(N_S) = \sum_{i=0}^{n} f_n^{(i)} \ln^i(N_S)\,, \tag{5.11}$$

and abbreviate $f_n \equiv f_n^{(0)}$. For $i > 0$ the coefficients $f_n^{(i)}$ are fixed by the lower orders (coefficients f_m with $m < n$) and the coefficients β_j of the β-function for $j \leq n-1$.

The validity of the ansatz (5.10) can also be illustrated considering the perturbative expansion of the leading dressed gluon propagator $D(k) \propto 1/k^2$, where $k_4 = 0$ in the static limit and lattice corrections are neglected.[36] With the formal ultraviolet cutoff $1/a$ and an infrared cutoff $1/(aN_S)$, we may write

$$\begin{aligned}\delta m(N_S) &\propto \int_{1/(aN_S)}^{1/a} dk\, k^2 D(k) \\ &\sim \frac{1}{a}\sum_n c_n \alpha^{n+1}\left(a^{-1}\right) - \frac{1}{aN_S}\sum_n c_n \alpha^{n+1}\left[(aN_S)^{-1}\right]\,.\end{aligned} \tag{5.12}$$

When re-expressing $\alpha\left[(aN_S)^{-1}\right]$ in terms of $\alpha(a^{-1})$, we may consider two situations:
(a) $N_S > e^n$. On such big lattices the second term of Eq. (5.12) is exponentially suppressed in n and the renormalon is directly seen in the coefficients $c_n(N_S)$.
(b) $N_S < e^n$. The second term of Eq. (5.12) is important and the renormalon cancels order-by-order in n. It is easy to visualize the importance of this term in the large-β_0

[36] Compare with Sec. 3.7.1 and the discussion around Eq. (3.42).

limit [211], in which the asymptotic behaviors of c_n and $f_n^{(i)}(N_S)$ are expected to read

$$c_n \simeq N_m \left(\frac{\beta_0}{2\pi}\right)^n n!\,, \qquad f_n^{(i)}(N_S) \simeq N_m \left(\frac{\beta_0}{2\pi}\right)^n \frac{n!}{i!}\,. \tag{5.13}$$

Inserting the expression for $f_n^{(i)}(N_S)$ in Eq. (5.11), we see that the sum approaches an exponential such that we can write

$$\begin{aligned} c_n - \frac{f_n(N_S)}{N_S} &\simeq N_m \left(\frac{\beta_0}{2\pi}\right)^n n! \left[1 - \frac{1}{N_S}\left(e^{\ln N_s} - \frac{1}{N_S}\sum_{s=n+1}^{\infty}\frac{1}{s!}\ln^s N_S\right)\right] \\ &= \frac{1}{N_S} N_m \left(\frac{\beta_0}{2\pi}\right)^n n! \sum_{s=n+1}^{\infty}\frac{1}{s!}\ln^s(N_S)\,. \end{aligned} \tag{5.14}$$

One can easily verify that this infinite sum converges and vanishes $\propto \frac{1}{n!}$. Overall, the factorial in n and hence the renormalon cancel from Eq. (5.14): we are left with a convergent series of coefficients $c_n(N_S)$. This is exactly the situation ($N_S < e^n$) we find ourselves in with present-day numerical simulations, so in order to reveal the conjectured $d = 1$ renormalon we need to carefully treat the FSE. A similar phenomenon was numerically observed for the static singlet energy $E(r)$ we mentioned above [212, 213]. We will check in our fits whether the coefficients f_n that parametrize the $1/N_S$ corrections indeed yield $f_n \simeq c_n$ for large n and support the above renormalon dominance picture.

In the $N_T \to \infty$ limit and up to $\mathcal{O}(N_S^{-2})$ effects, the fitting function for $c_n(N_S)$ depends on c_n, f_m with $m \leq n$, and the beta coefficients β_j with $j < n$. If we work in the lattice scheme, the unknown β_3 and higher (see Appendix A.1) will start to play a role at $\mathcal{O}(\alpha^5)$. One may try to introduce them as free parameters in our fit but our data is not sensitive enough to determine them. This entails an uncertainty which we have to consider in the error analysis, at the same time the robustness of our fits despite the unknown β_j is comforting and has a profound physical reason. Their impact diminishes with increasing order n as the behavior of the coefficients f_n starts to be governed by the $d = 1$ renormalon.[37] Preceeding coefficients f_m with $m < n$ would then be suppressed and so would the higher β_j that multiply them. Let us try to quantify the suppression of lower f_m in a large-n analysis. In the large-β_0 limit and assuming renormalon dominance for the coefficients f_n, we find [cf. Eqs. (5.13) and (5.14)]

$$f_n(N_S) = N_m \left(\frac{\beta_0}{2\pi}\right)^n n! \sum_{s=0}^{n}\frac{1}{s!}\ln^s(N_S)\,. \tag{5.15}$$

Note that terms with higher powers of $\ln(N_S)$ (with $s \lesssim n$) are suppressed by a term $\sim 1/n!$. Therefore, at sufficiently large orders n the factorial exceeds the large logarithms. This suppression also holds beyond the large-β_0 limit. Irrespectively of that, our knowledge of

[37] It is common belief that the coefficients of the β-function in the $\overline{\text{MS}}$ scheme are convergent or at least diverge slower than factorially (see the related discussion in Sec. 3.4 of [37]).

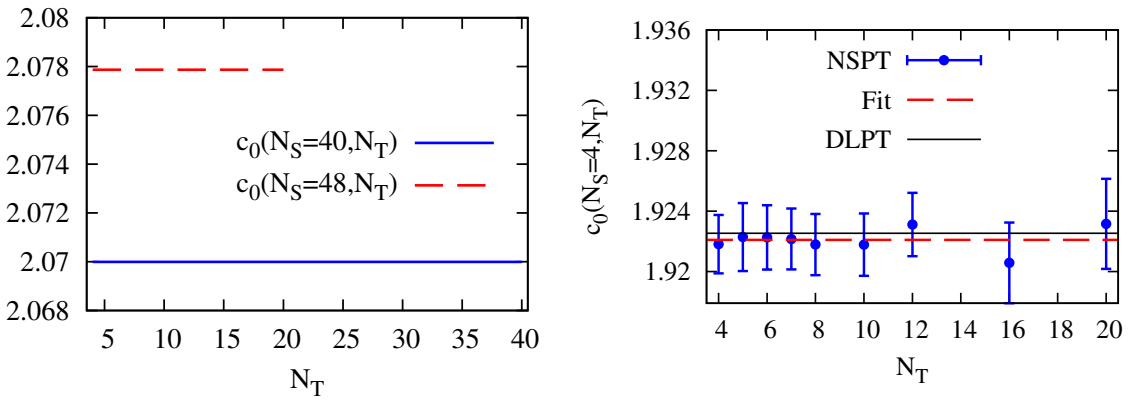

Fig. 5.3: The one-loop self-energy from DLPT with PBC (left) and from NSPT with TBC (right). DLPT results are shown for $40^3 \times N_T$ and $48^3 \times N_T$ lattices volumes. NSPT data are displayed for $4^3 \times N_T$ lattice volumes, together with the result for fitting a constant and the DLPT prediction Eq. (5.18) for $N_S = 4$.

the coefficients $\beta_0, \beta_1, \beta_2$ allows us to pin down $f_n^{(n),(n-1),(n-2)}$, the contributions of the lowest coefficients $f_{0,1,2}$ at a given order n. The question remains whether terms $\ln^s(N_S)$ with $s \sim 1$, which have no factorial suppression, cause problems. The successive inclusion of β_j with $j > 0$ follows a hierarchic structure: The running associated with β_1 produces $1/n$ suppressed terms to Eq. (5.13), β_2 produces $1/n^2$ suppressed terms and so on. For instance, $f_n^{(1)}$ would yield

$$f_n^{(1)} = N_m \, \nu \left(\frac{\beta_0}{2\pi}\right)^n \frac{\Gamma(n+1+b)}{\Gamma(1+b)} \left[1 + \frac{b}{(n+b)} \frac{n-1}{n-1+b} + \mathcal{O}(\beta_2/n^2)\right] . \tag{5.16}$$

The claim that our ignorance of higher β_j (with $j > 2$ in the lattice scheme) is not decisive and hence the truncation after β_2 is justified will be put on test with three different fits including β_j for $j \leq 0, 1$ and 2.

5.1.4 Finite-N_T effects

We now try to assess the effect of finite temporal lattice extents N_T on the c_n. In addition to our own data, we now have at our disposal precise DLPT results for the unsmeared triplet coefficients $c_{0,1}^{(3,0)}$ at one and two loops, which were not available at the time of writing [204]. Using the prescription given in [44], the DLPT data were obtained for $n = 0$ with TBCxyz and for $n = 0, 1$ with PBC. The prescription is correct for $n = 0$ but difficult to implement for $n = 1$ for $N_T \gtrsim N_S$. We will therefore exclude this range from our analysis and focus on small $N_T \leq N_S/2$ values, which are the most interesting for the current discussion anyhow.

In the left-hand plot of Fig. 5.3, we display the leading-order DLPT self-energy coefficients using PBC for $N_S = 40$ and $N_S = 48$ and various values of N_T. Error bars can be safely omitted due to the high numerical precision of the diagrammatic approach. We see

that at leading order there is no dependence on N_T,

$$c_0(N_S, N_T) = c_0(N_S)\,. \tag{5.17}$$

More specifically, with DLPT and TBCxyz one finds

$$c_0(N_S) = c_0 - \frac{f_0}{N_S} - \frac{w_0}{N_S^3} + \mathcal{O}\left(\frac{1}{N_S^4}\right)\,, \tag{5.18}$$

where the existence of a $\mathcal{O}(N_S^{-2})$ term can be ruled out at leading order and the coefficients read

$$\begin{aligned} c_0 &= 2.1172743570834807985970\ldots\,, \\ f_0 &= 0.76962563284(2)\,, \\ w_0 &= 0.14932(3)\,. \end{aligned} \tag{5.19}$$

Our value for c_0 (and c_1 below) is an update [204] of the previous direct [118, 119, 214] and indirect[38] determinations [44, 215] within DLPT.

The behavior described by Eq. (5.18) is also seen in our TBCxyz data, as the right-hand plot of Fig. 5.3 illustrates for NSPT results from $4^3 \times N_T$ volumes. Fitting a constant to $c_0(N_S = 4, N_T)$ NSPT data yields $c_0(N_S = 4) = 1.9221(7)$, in perfect agreement with $c_0(N_S = 4) = 1.92253$, the diagrammatic prediction from Eq. (5.18) evaluated at $N_S = 4$.

The situation changes at two loops. For DLPT at finite N_S and N_T, we only have results using PBC. They are shown in the left-hand plot of Fig. 5.4. We change the horizontal axis to N_T^{-3} and obtain straight lines for $c_1(40, N_T)$ and $c_1(48, N_T)$. In very good approximation, it seems to hold

$$c_1(N_S, N_T) = c_1(N_S)\left[1 + \frac{k_1 + k_2 N_S}{N_T^3} + \mathcal{O}\left(\frac{1}{N_T^4}\right)\right]\,, \tag{5.20}$$

with constants k_1 and k_2. We find for DLPT and PBC in the infinite-volume limit

$$c_1 = 11.1425(25)\,. \tag{5.21}$$

Since this is valid regardless of the choice of boundary conditions, we will later on take Eq. (5.21) as a reference for our TBCxyz data.

We compare the parametrization Eq. (5.20) with our results using NSPT and TBCxyz. It turns out that a two-parameter fit with this ansatz for fixed $N_S = 4$ works perfectly with $\chi^2_{\text{red}} = 0.996$ (right-hand plot of Fig. 5.4). Eq. (5.20) indeed seems to capture the N_T dependence and we expect a similar behavior at higher orders. A faithful implementation of Eq. (5.20) in a fit has the disadvantage that NSPT data do not have the required precision to cope with that many fit parameters. On the contrary, such a strategy would inappro-

[38] One can also use the perimeter law for large Wilson loops.

priately inflate the error on those parameters which we can extract with good accuracy, such as the infinite-volume coefficients c_n.

In other words, we need a simplified ansatz that reflects the precision of our data and yet minimizes the uncertainty due to N_T. Most of our geometries yield $N_T \gtrsim N_S$. Hence, we estimate that terms $\propto N_S N_T^{-3} \approx N_T^{-2}$ are subleading compared to the N_S^{-1} correction we found in Sec. 5.1.3. Then, one solution is to exclude all $c_n(N_S, N_T)$ with $N_T \leq N_T^{\min}$ from the analysis. To be on the safe side, we additionally enforce $N_T \geq N_S$.

At fixed N_S we find that for $N_T^{\min} = 10$ the $\mathcal{O}(N_T^{-3})$ corrections have diminished to a point where they may be safely neglected. This is illustrated in the right-hand plot of Fig. 5.4: discarded data points are shown with open symbols, whereas filled symbols mark the data points which fulfill $N_T > N_T^{\min}$. As can be seen in this region, the N_T dependence vanishes and $c_1(N_S = 4, N_T > N_T^{\min}) \sim c_1(N_S = 4)$ within errors.

We will perform four alternative fits, two using $N_T^{\min} = 8$ or 10 and two with smaller $N_T^{\min} = 4$. The augmented N_T effects of the latter shall be counter-balanced by two different effective parametrizations of the N_T dependence. They are summarized in the last term of

$$c_n(N_S, N_T) = c_n - \frac{f_n(N_S)}{N_S} + h_{\text{eff}}(N_S, N_T)\,. \tag{5.22}$$

Eq. (5.20) tells us that we should allow for a N_S dependence of the N_T correction. For instance, we can introduce a dependence on the running coupling in analogy to $f_n(N_S)$. We do so by setting

$$h_{\text{eff}}^{(a)}(N_S, N_T) = \frac{v_n(N_S)}{N_T^2}\,, \tag{5.23}$$

with one additional fit parameter v_n per order n. We can just as well employ

$$h_{\text{eff}}^{(b)}(N_S, N_T) = \frac{\tilde{v}_n N_S}{N_T^3}\,, \tag{5.24}$$

where $\tilde{v}_n$ also means one additional fit parameter per order n.

The effective parametrizations $h_{\text{eff}}^{(a),(b)}$ have been chosen based on χ^2_{red} and how well they reproduce the infinite-volume coefficient c_1. The precise outcome for the coefficients $v_n, \tilde{v}_n$, however, is irrelevant.[39] What matters is that the main fit result, the infinite-volume coefficients c_n, remains stable regardless of whether one works with $N_T^{\min} = 10$ or 8 or $h_{\text{eff}}^{(a),(b)}$.

5.1.5 Fits

The following fitting strategy emerges from the considerations of Secs. 5.1.4 and 5.1.3: we attempt a global fit with three independent variables N_S, N_T, n and two new fit parameters c_n, f_n per order n. Our parametrization Eq. (5.9) is quite efficient, given that

[39] In that sense, what we applied in [204] amounts to just another effective parametrization $h_{\text{eff}}^{(c)}$.

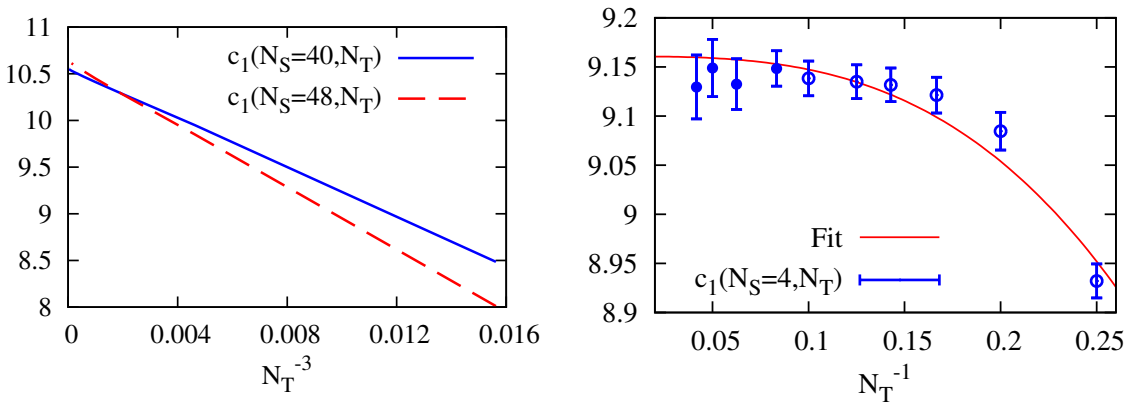

Fig. 5.4: The two-loop self-energy from DLPT with PBC (left) and from NSPT with TBC (right). DLPT results are shown for $40^3 \times N_T, 48^3 \times N_T$ lattices volumes. NSPT data is displayed for $4^3 \times N_T$ lattice volumes together with the fit result using Eq. (5.20).

19 lattices volumes fulfill the constraint $N_T > \min(10, N_S)$ at two loops and still four volumes at 20 loops (cf. Table 5.1). For the variants $N_T^{min} = 8$ or $h_{\text{eff}}^{(a),(b)}$ there are even more degrees of freedom. We exploit the leading-order DLPT results by fixing the values for c_0 and f_0 according to Eq. (5.19), i.e., the global fit starts at $n = 1$.[40] We do not constrain the value for c_1 as we want to use the DLPT value [Eq. (5.21)] as a benchmark.

In a first step, we conduct three different fits including β_j for $j \leq 0, 1$ and 2. The outcome for χ^2_{red} and the c_n is listed in Table A.1 of Appendix A.5 (as are all following tables). The goodness of fit is approximately equal in all cases and yields a good $\chi^2_{\text{red}} \approx 1.264$ for the global fit using $\{\beta_0, \beta_1, \beta_2\}$.

To visualize the quality of this fit for $n = 6$ and 9, we contrast in Figs. 5.5 and 5.6 the data points that were taken into consideration with the fitting function evaluated at the central values for the parameters c_n and f_m with $m \leq n$. From lowest to largest N_S, the finite-volume coefficients are reproduced very well by our fit function. The $n = 9$ infinite-volume result, for comparison, yields $c_9 \sim 3.75 \times 10^9$ and lies significantly above the range covered by the data points, indicating sizable N_S^{-1} corrections at large orders. They grow systematically with n: the corrections are less pronounced at $n = 6$ than at $n = 9$, causing that the various curves lie closer together in Fig. 5.5. The above findings clearly support the theoretical picture which emerged in Sec. 5.1.3.

Two things are worth noting regarding our limited knowledge of the β-function:
(a) The $\{\beta_0\}$ fit is enough to describe the data to an acceptable degree.
(b) It is evident from the resulting c_n that the change from the $\{\beta_0\}$ fit to $\{\beta_0, \beta_1\}$ is greater than that from $\{\beta_0, \beta_1\}$ to $\{\beta_0, \beta_1, \beta_2\}$, in support of the convergence behavior we conjectured around Eq. (5.16). From both we conclude that the error due to the unknown higher β_i is under control. As an estimate for this theoretical uncertainty σ_β, we take the relative difference between the central values of the $\{\beta_0, \beta_1, \beta_2\}$ and $\{\beta_0, \beta_1\}$ fits.

[40] As f_0 will propagate to all orders of the fitting function via $f_n^{(n)}$, it is convenient (but not necessary) to fix it to the precise DLPT value instead of fitting it from NSPT data.

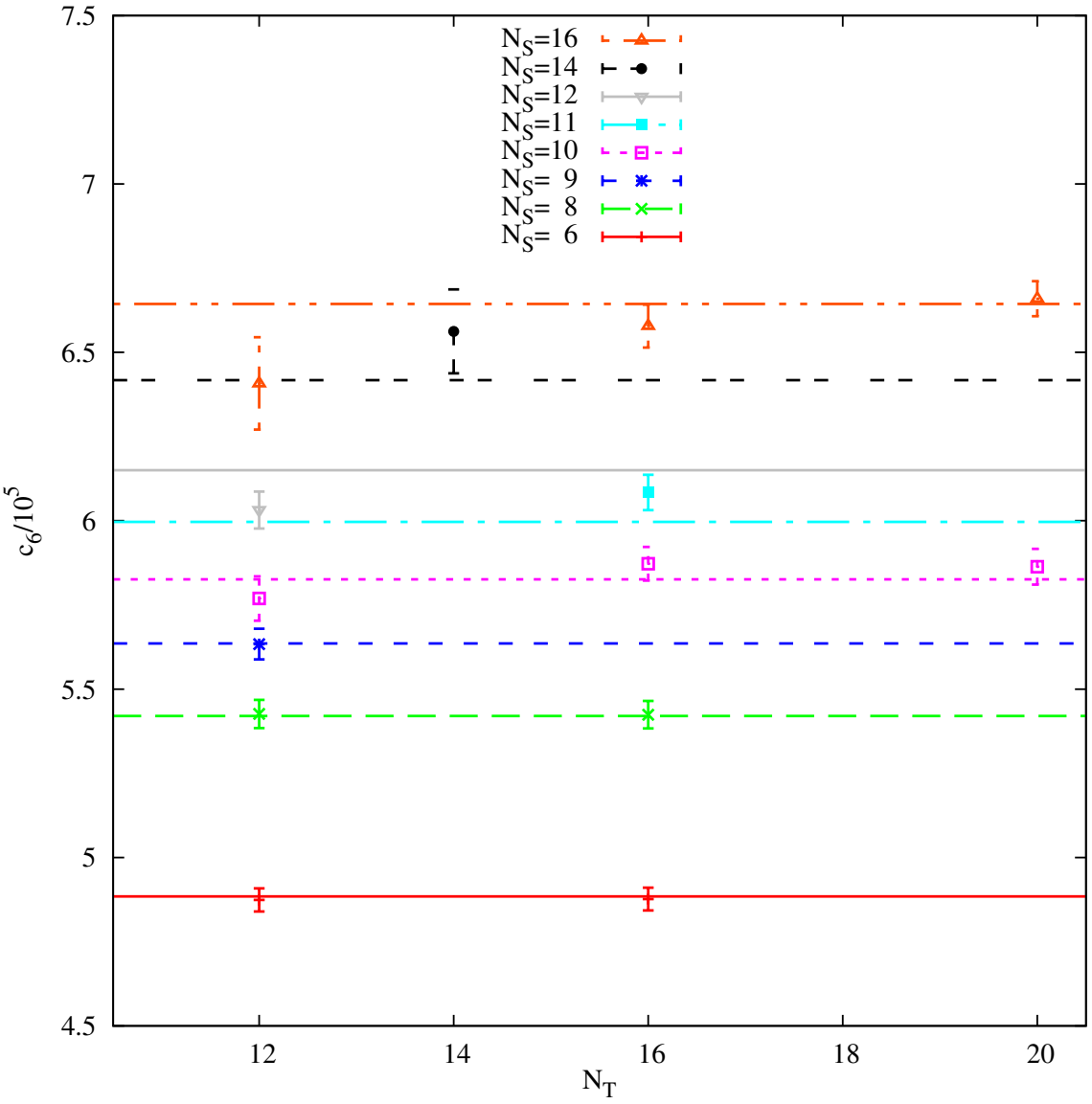

Fig. 5.5: Comparison between the global fit result using $\{\beta_0, \beta_1, \beta_2\}$ and the simulation data for $n = 6$. Shown are only those data points with $N_T \geq 10$ and $N_T \geq N_S$ that have been considered in the fit.

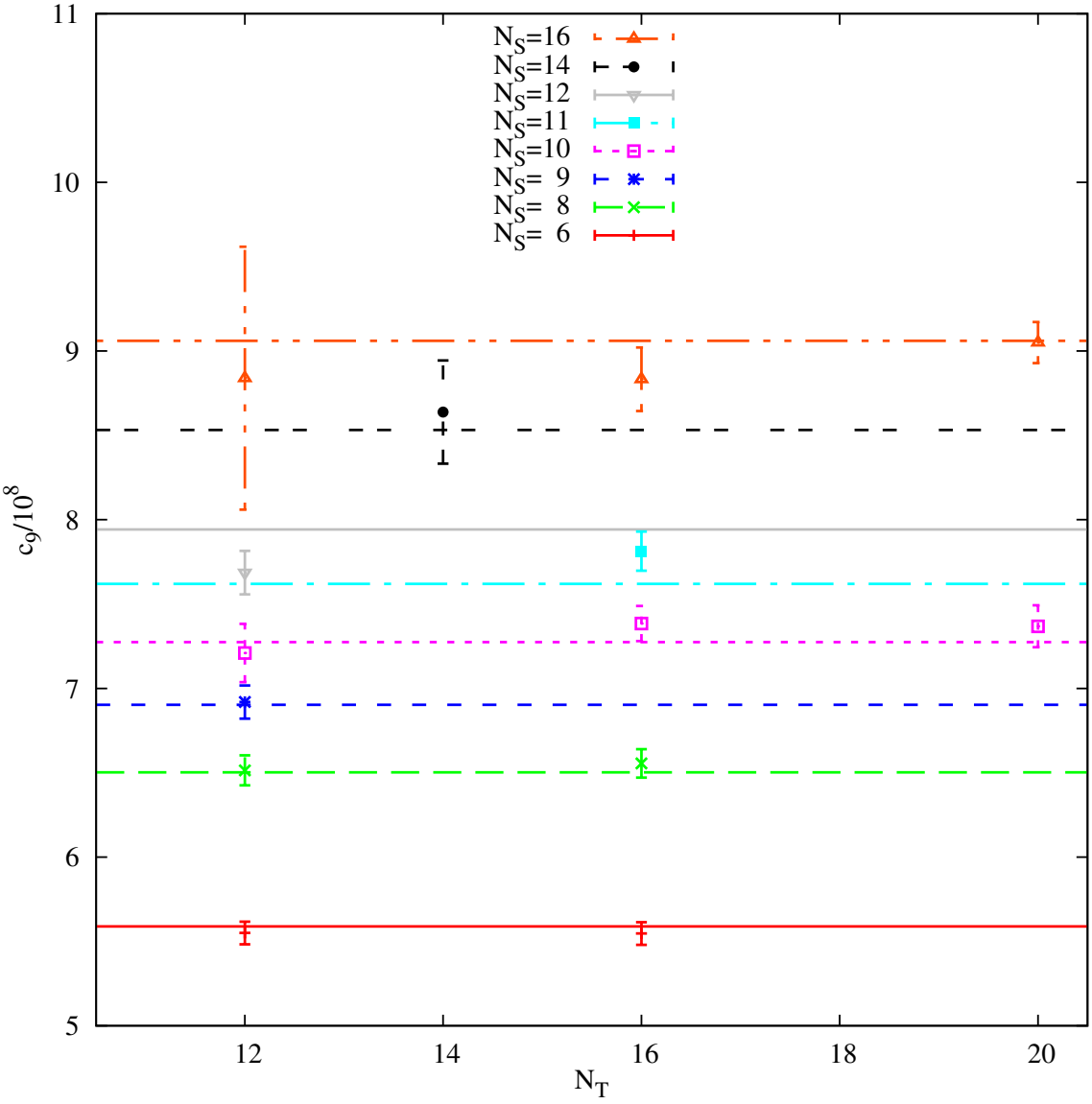

Fig. 5.6: Comparison between the global fit result using $\{\beta_0, \beta_1, \beta_2\}$ and the simulation data for $n = 9$. Shown are only those data points with $N_T \geq 10$ and $N_T \geq N_S$ that have been considered in the fit.

In order to assess the discriminative power of our data and appreciate the χ^2_{red} value reported above, we perform a simple test. We do a fit with the above setup, the only difference is that we fit from lowest order and without running in the N_S^{-1} correction. We obtain a poor $\chi^2_{\text{red}} \approx 2.868$ and the outcome for c_1 deviates from the DLPT value Eq. (5.21) by $\sim 20\ \sigma$. The bad result has nothing to do with the fact that we start fitting at $n = 0$ and do not make use of DLPT: we achieve an excellent $\chi^2_{\text{red}} \approx 1.152$ and reproduce the literature value for c_1 within errors with this very setup as soon as we switch the running on again. We infer that the incorporation of the renormalization group running for α in the parameters $f_n(N_S)$ is indispensable.

We turn to the finite-N_T effects that have been thoroughly discussed in Sec. 5.1.4. The outcome for the variants $N_T^{\min} = 8$ and $h_{\text{eff}}^{(a),(b)}$ is shown in Table A.2. It should be compared with the corresponding $N_T^{\min} = 10$ result in the $\{\beta_0, \beta_1, \beta_2\}$ column of Table A.1. It is conceivable that a lowering of $N_T^{\min}$ from 10 to 8 leads to an increase in χ^2_{red}. The additional fit parameters due to an incorporation of $h_{\text{eff}}^{(a),(b)}$ more than compensate a small $N_T^{\min} = 4$, resulting in lower χ^2_{red}. What really matters, though, is that the infinite-volume coefficients c_n vary little regardless of the N_T prescription we choose. The alternatives listed in Table A.2 work equally well, so we pick the difference between the central values of $N_T^{\min} = 10$ and 8 as an estimate for the uncertainty σ_T we incur by omitting the strict $N_T \to \infty$ limit.

The outcome is different if we make an analogous variation of the minimal values $N_S^{\min}$ that we allow for the spatial extent. The resulting χ^2_{red}, c_1 and c_{19} are shown in Table A.3 for fixed $N_T^{\min} = 10$ and $N_T \geq N_S$. The goodness of fit is best if we include all our geometries (this corresponds to $N_S^{\min} = 3$) instead of eliminating some of them by increasing $N_S^{\min}$. This inflates the error on the coefficients c_n, and the central values for c_1 move away from the DLPT result Eq. (5.21). Overall, we see no reason for a constraint $N_S^{\min} > 3$ since the N_S^{-1} parametrization apparently works well without.

We have completed the exploration of possible sources of error and list the final numbers for all $c_n^{(R,\rho)}$ in Table A.4. The quoted errors result from summing statistical and theoretical uncertainties in quadrature. Schematically, we have at each order n

$$\sigma_{\text{final}} = \sqrt{\sigma^2_{\text{stat.}} + \sigma^2_\beta + \sigma^2_T}\,. \tag{5.25}$$

Note that we find $\sigma_\beta \gg \sigma_T, \sigma_{\text{stat.}}$ on average, i.e., the unknown higher β_j are the dominant source of error.

So far we have always relied on c_1 as a benchmark. However, our three-loop result $c_2^{(3,0)} = 86.10(13)$ also has antecessors in the literature, namely, 86.2(6) from large Wilson loops in NSPT [216] and 86.6(5) from Monte Carlo simulations at weak coupling [197]. We see that the agreement is excellent.

As far as smeared and octet data are concerned, we learn two things from Table A.4:
(a) The coefficients $c_n^{(R,0)}$ and $c_n^{(R,1/6)}$ differ substantially only at small n, so smearing

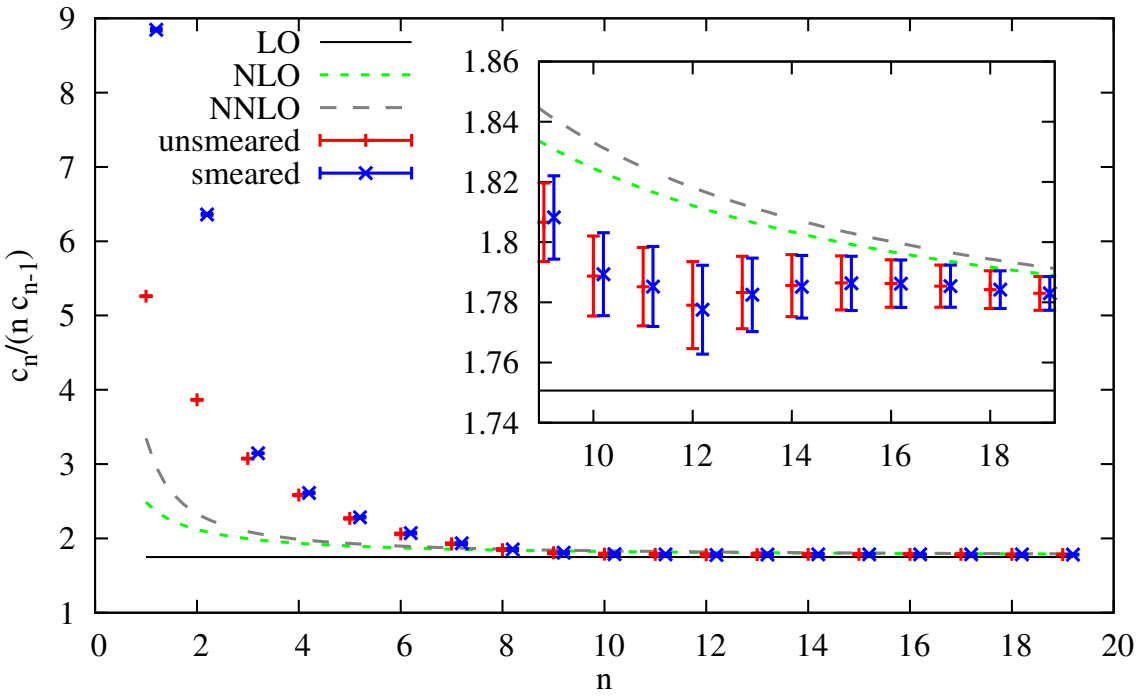

Fig. 5.7: Ratios $c_n/(nc_{n-1})$ of the smeared (blue) and unsmeared (red) triplet static self-energy coefficients c_n in comparison to the theoretical prediction Eq. (5.5) at different orders in the $1/n$ expansion.

effects are indeed confined to low orders, as we suspected in Sec. 5.1.1.
(b) We list octet data $c_n^{(8,\rho)}$ rescaled with an additional factor of C_A/C_F and realize that the values more or less equal the corresponding $c_n^{(3,\rho)}$ from the triplet. We know already that the self-energy in the fundamental representation carries the corresponding color factor C_F [see Eqs. (3.41) and (3.67)], its generalization to C_R in an arbitrary representation R seems logical. In a more general context, it goes under the name of *Casimir Scaling* of static QCD potentials [217]. In the case of the perturbative static potential, violations are found at $\mathcal{O}\left(\alpha^4\right)$ [218]. The most probable explanation why we do not see any is that our data are not sensitive enough to resolve the $\mathcal{O}(N_c^{-2})$ violations (N_c being the number of colors).

We now turn to the discussion of ratios $c_n/(nc_{n-1})$. This secondary observable draws its importance from Eq. (5.5) and the fact that it provides a graphic test whether we see renormalon dominance in our coefficients or not. Although we extract them from the coefficient fit result, ratios can be determined with smaller statistical errors $\sigma_{\text{stat.}}$ than those of the c_n themselves, in particular at high orders n. The reason is the strong positive correlation between fit parameters c_{n-1} and c_n, which at large n are increasingly determined by the input for the renormalization group running. Small errors on ratios are familiar from the underlying NSPT simulations, albeit for a different reason: fluctuations in low-order coefficients induce similar fluctuations at higher orders and therefore correlate in a positive way. It is possible then that ratios of two subsequent coefficients vary very little. We find $\sigma_\beta, \sigma_T \gg \sigma_{\text{stat.}}$ for the ratios and again we apply Eq. (5.25) to arrive at an overall error.

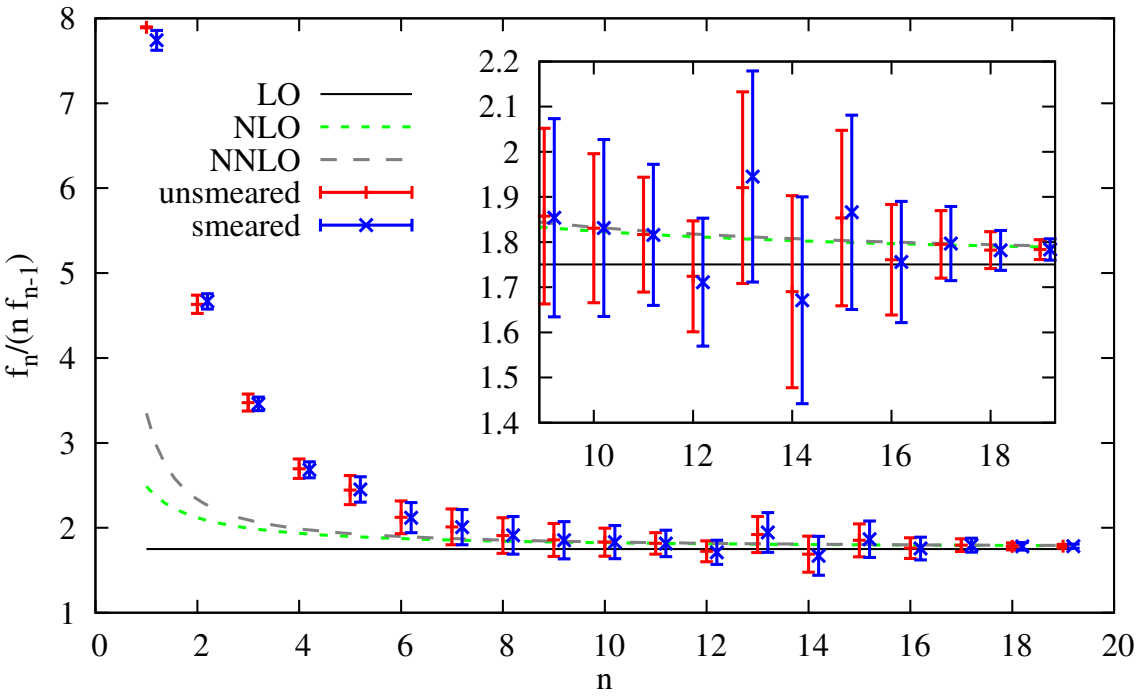

Fig. 5.8: Ratios $f_n/(nf_{n-1})$ of the N_S^{-1} corrections f_n of the smeared (blue) and unsmeared (red) triplet static self-energy in comparison to the theoretical prediction Eq. (5.5) at different orders in the $1/n$ expansion.

In Fig. 5.7, we show our ratios from smeared (blue) and unsmeared (red) triplet data, together with the constant leading-order ratio prediction[41] (LO) and the two corrections for β_1 (NLO) and β_2 (NNLO) according to the $1/n$ expansion Eq. (5.5). We confirm that our data yield the prediction of a leading $d = 1$ infrared renormalon in the static self-energy. Smeared data behave differently at low orders due to a different discretization but has the same long-distance behavior as dictated by the infrared renormalon. The onset of renormalon dominance is located at $n \approx 9$. At orders $n \gtrsim 9$, the behavior of coefficients c_n is fully determined by the asymptotic formula Eq. (5.3). As the zoom inlay of large orders in Fig. 5.7 reveals, the ratio result is even precise enough to distinguish between LO on the one side and $1/n$ corrections on the other side. We can also confirm with Fig. 5.8 our claim in Sec. 5.1.3 that the coefficients f_n which parametrize the N_S^{-1} correction due to the replica potential exhibit the same leading $d = 1$ renormalon. It is evident from Table A.4 and the discussion above that the outcome is identical for octet coefficients $c_n^{(8,\rho)}$.

Since we not only see the renormalon but also went well beyond (~ 10 orders) the onset of renormalon dominance, it should be possible to extract the normalizations $N_{m,A}$ for the static self-energies of fundamental and adjoint sources. As they cancel from the ratio computation, we revert to the coefficients c_n and use Eq. (5.3) to solve for the normalization. The outcome is shown as a function of n in Fig. 5.9 for unsmeared fundamental sources. It is complementary to our previous findings: in the renormalon regime, the normalization forms a plateau, $N_m\,(n \gtrsim 9) \sim N_m$. At the highest orders, the dependence on n vanishes

[41] It corresponds to the large-β_0 approximation.

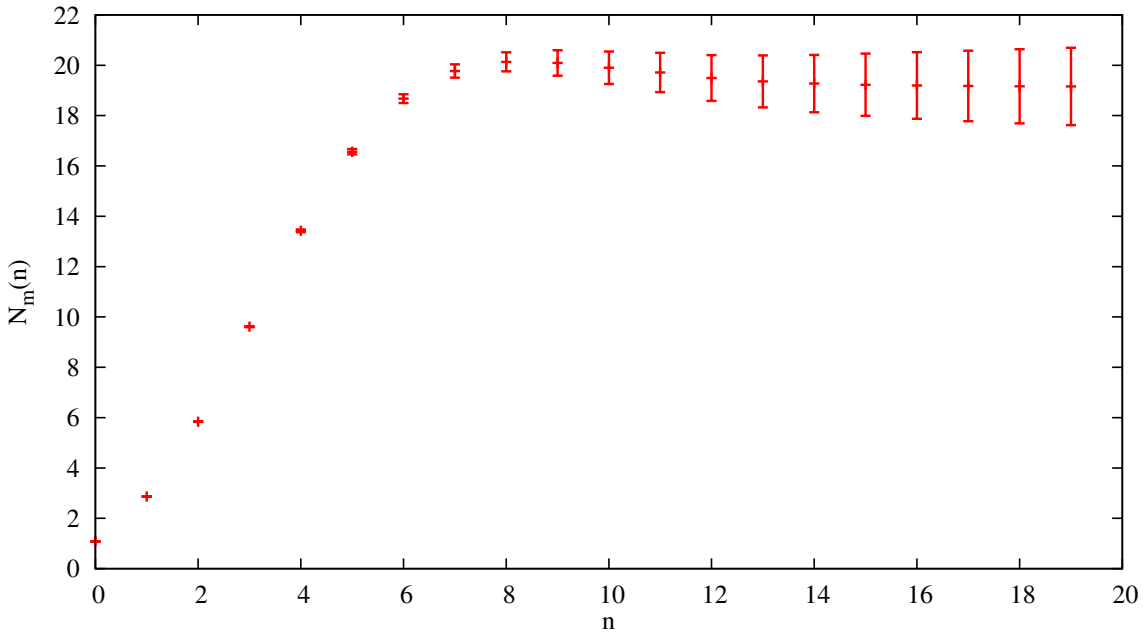

Fig. 5.9: The normalization N_m obtained from Eq. (5.4) as a function of the order n for the unsmeared triplet coefficients $c_n^{(3,0)}$.

almost completely, and we therefore take $N_m(n = 19)$ as our final value. In addition, we exploit the fact that smeared and unsmeared data have the same behavior at large n and hence the same normalization. Averaging both numbers and taking the maximal error we obtain

$$N_m^{\mathrm{L}} = 19.1(1.6)\,, \qquad \frac{C_F}{C_A} N_A^{\mathrm{L}} = 18.8(1.8)\,. \tag{5.26}$$

The superscript indicates the lattice scheme.

The conversion to the more standard $\overline{\mathrm{MS}}$ scheme is exact and determined by the first coefficient d_1 of the scheme conversion from lattice to $\overline{\mathrm{MS}}$ [Eq. (A.8) in Appendix A.1]:

$$N_{m,\Lambda}^{\mathrm{L}} = N_{m,A}^{\overline{\mathrm{MS}}}\, e^{\frac{2\pi d_1}{\beta_0}}\,. \tag{5.27}$$

The exponential yields the ratio of Λ_{QCD} evaluated in the two schemes [Eq. (A.10)]. In this light, Eq. (5.27) can be understood since the ambiguity of the perturbative series due to the renormalon is universal: there is an uncertainty $\sim N^X \Lambda_{\mathrm{QCD}}^X$ inherent to any determination of m_{pole}, regardless of the scheme X one chooses.[42] We use the value for $\Lambda_{n_f=0}^{\overline{\mathrm{MS}}}$ given in [219] and obtain

$$N_m^{\overline{\mathrm{MS}}} \Lambda_{n_f=0}^{\overline{\mathrm{MS}}} = 153(13)(13)\ \mathrm{MeV}\,, \qquad N_A^{\overline{\mathrm{MS}}} \Lambda_{n_f=0}^{\overline{\mathrm{MS}}} = 151(14)(12)\ \mathrm{MeV}\,. \tag{5.28}$$

The errors in brackets refer to the errors in $N_{m,A}$ and $\Lambda_{n_f=0}^{\overline{\mathrm{MS}}}$, respectively. The ambiguity is in the range of the literature values quoted around Eq. (3.48) in Sec. 3.7.1. The

[42] See also the discussion around Eq. (3.48) in Sec. 3.7.1.

normalizations $N_{m,A}$ in the $\overline{\text{MS}}$ scheme read

$$N_m^{\overline{\text{MS}}} = 0.662(57)\,, \qquad \frac{C_F}{C_A} N_A^{\overline{\text{MS}}} = 0.652(62)\,. \tag{5.29}$$

Although our approach is completely different and the separation of scales between lattice and $\overline{\text{MS}}$ is large [Eq. (A.10)], our results in Eq. (5.29) are in very good agreement with direct determinations in the continuum, e.g., $N_m^{\overline{\text{MS}}} \simeq 0.62$ [212].

As opposed to the perturbative ambiguity, the onset of renormalon dominance at a given order n is scheme-dependent. The continuum computations mentioned above rely on the assumption that the $\overline{\text{MS}}$ scheme is ideal for renormalon studies[43] because the renormalon dominance should be seen at significantly lower orders n in the perturbative expansion than in the lattice scheme, for instance. We can try to verify this claim by mimicking a scheme conversion of our lattice data to $\overline{\text{MS}}$.

We can do so at a very early stage in the data analysis [210] since it allows us to directly evaluate secondary observables on the measurements of the primary observable, in our case the Polyakov loop.[44] However, the scheme change from lattice to $\overline{\text{MS}}$ is known only up to the three-loop coefficient d_2 [Eq. (A.8)], so we can only approximatively convert our $\mathcal{O}(\alpha^{20})$ data. There is more than one way to do this, and what we find works best is to rewrite Eq. (A.8) as

$$\alpha_{\text{L}}(\mu) = \alpha_{\overline{\text{MS}}}(\mu) \left[1 + d_1 \alpha_{\overline{\text{MS}}}(\mu) + \left(d_2 - d_1^2\right) \alpha_{\overline{\text{MS}}}^2(\mu)\right]^{-1} . \tag{5.30}$$

If the expansion is done in this way, we generate a lot more terms in high orders. The question then is whether the approximation (which defines a scheme of its own) is enough to maintain the main characteristics of the true $\overline{\text{MS}}$ scheme. It apparently is the case, as Fig. 5.10 shows for a $\{\beta_0, \beta_1, \beta_2^{\overline{\text{MS}}}\}$ fit to the fake $\overline{\text{MS}}$ data for the unsmeared fundamental self-energy, starting at $n = 0$. First of all, note the negative ratio values at low n, indicating c_n with alternating sign. Furthermore, we see a rapid convergence to the leading $d = 1$ renormalon expectation around $n \sim 5$, so indeed a lot earlier than in the lattice scheme (cf. Fig. 5.7). Similarly, the theoretical prediction for ratios at different orders in the $1/n$ expansion shows a faster convergence in the fake $\overline{\text{MS}}$ scheme, the NLO and NNLO curves collapse onto each other at small n.

Since the approximate $\overline{\text{MS}}$ conversion yields promising results, we compute the normalization directly in this approach. The error entailed by the approximate conversion Eq. (5.30) is difficult to estimate, but for the central value we get from the above fit $N_m^{\overline{\text{MS}}} \simeq 0.53$. In spite of the incomplete scheme conversion, this agrees reasonably well with our main result Eq. (5.29) for the $\overline{\text{MS}}$ scheme.

[43] See [206, 212, 213], for instance.

[44] Its logarithm is the most obvious secondary observable and produces the coefficients c_n. But we can also intertwine the logarithm with other functions such as a scheme change from lattice to $\overline{\text{MS}}$.

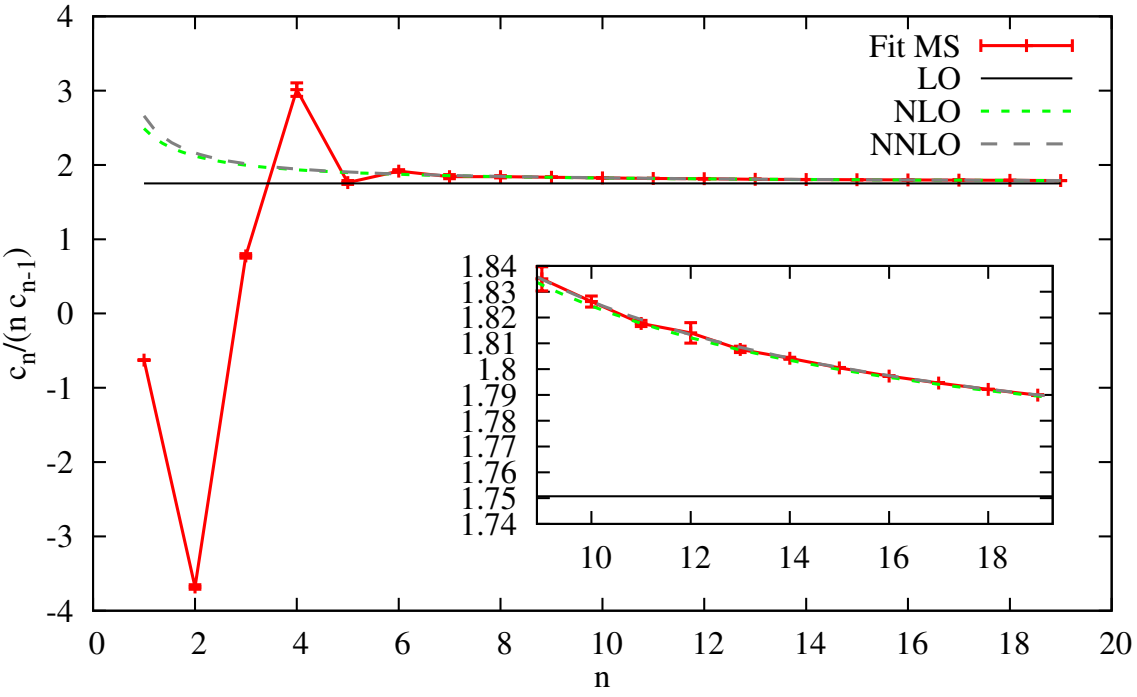

Fig. 5.10: Ratios $c_n/(nc_{n-1})$ of the unsmeared triplet static self-energy coefficients c_n from an approximate conversion to the $\overline{\text{MS}}$ scheme, Eq. (5.30). In comparison, the theoretical prediction Eq. (5.5) is shown at different orders in the $1/n$ expansion.

In our introduction to renormalons in Chap. 3 we came across the question regarding the order n_0 where the minimal term occurs in the perturbative expansion, indicating that the series should not be pursued beyond n_0. Now we are in the position to answer this question. We take our final set of self-energy coefficients for the unsmeared triplet and compute individual contributions $c_n\alpha^{n+1}$ to the resummed series. We do so for $\alpha = 0.07$ and 0.08. This choice is conceivable for the weak-coupling limit. We show the individual contributions in Fig. 5.11, together with the leading-order theoretical prediction Eq. (3.36) for n_0 in the case of a leading $d = 1$ IR renormalon. For both values of α the series must be truncated at fairly low orders n. It is a reflection of the fact that we have chosen an observable with a particulary strong renormalon ambiguity in the very beginning. We can also compute successive contributions $[a(\alpha)]^{-1}c_n\alpha^{n+1}$ in physical units (Fig. 5.12).[45] Interestingly, the sizes of the minimal term vary little – they lie in the range of 50–65 MeV.

This ends our discussion of the static self-energy. In Sec. 5.2 we consider an observable in which the renormalon ambiguity is much less pronounced, the elementary plaquette.

[45] The inverse lattice spacing turns the static self-energy into a dimensionful quantity [Eq. (4.3)]. We will discuss how one can obtain the lattice spacing $a = a(\alpha)$ around Eq. (5.40).

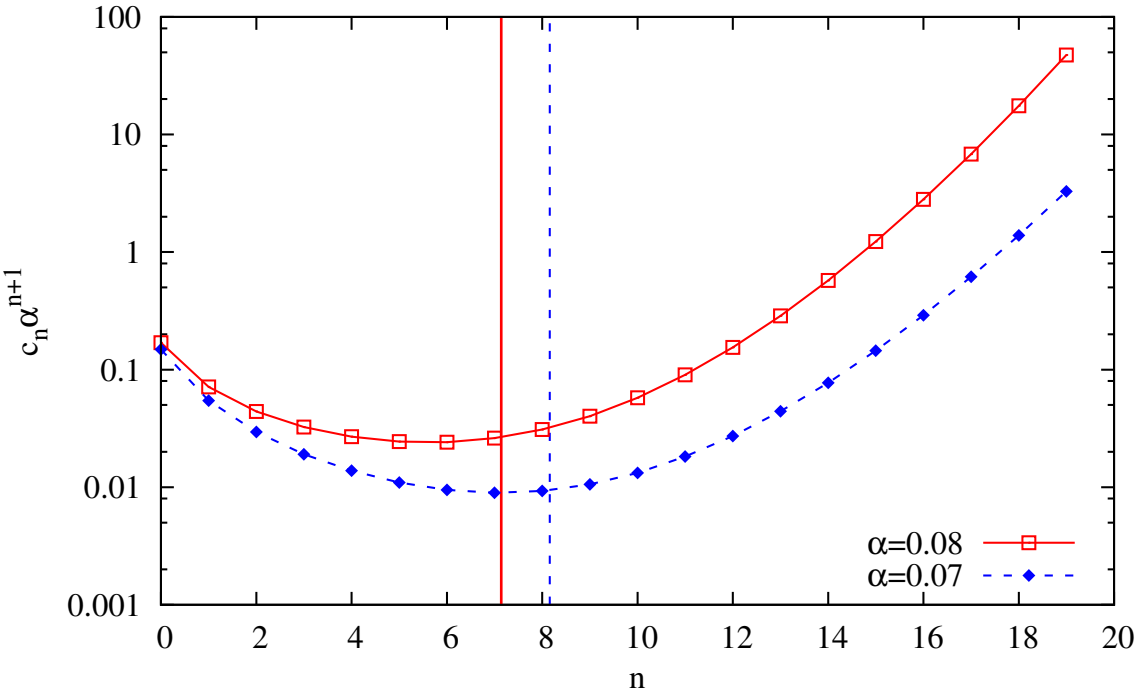

Fig. 5.11: Successive contributions $c_n\alpha^{n+1}$ to the resummed static self-energy for $\alpha = 0.08$ (red, solid) and $\alpha = 0.07$ (blue, dashed). Vertical lines in alike colors indicate the LO theoretical expectation for the order n_0 at which the minimal term $c_{n_0}\alpha^{n_0+1}$ should occur.

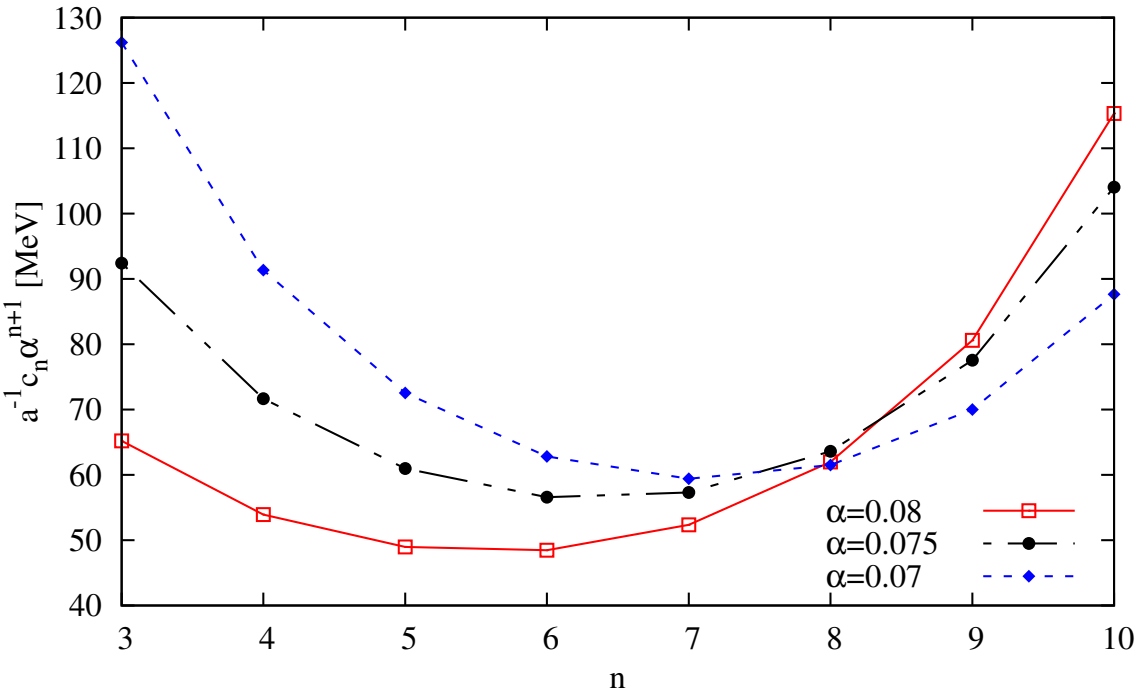

Fig. 5.12: Successive contributions $[a(\alpha)]^{-1}c_n\alpha^{n+1}$ in physical units for different values of the coupling α. Note that the minimal terms all lie in the range of 50–65 MeV.

5.2 The plaquette

Let us apply the experience we gained with the static energy to a different observable, the elementary plaquette. We introduced it already as the basic ingredient of the Wilson gauge action in Sec. 1.4, and considered it anew in Sec. 3.9.2 when discussing renormalons in theories with a hard regulator, like the lattice.

5.2.1 Prerequisites

Given that our analysis of the static energy in Sec. 5.1 was in line with the renormalon theory of Chap. 3, we assume that plaquette data should follow the prescription of Secs. 3.4 or 3.9.2, specifically: we search for an IR renormalon in an observable of dimension $d = 4$.

We can then take the asymptotic formulae for the static energy as in Sec. 5.1 and adopt it for plaquette purposes by proper insertions of d and the replacement

$$b \to b_p = db\,. \tag{5.31}$$

In the case of the plaquette, s_1 and higher have not been calculated. The coefficients yield

$$p_n \overset{n\to\infty}{=} N_p \left(\frac{\beta_0}{2\pi d}\right)^n \frac{\Gamma(n+1+b_p)}{\Gamma(1+b_p)} \left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right], \tag{5.32}$$

while ratios follow from

$$\frac{p_n}{p_{n-1}}\frac{1}{n} = \frac{\beta_0}{2\pi d}\left[1 + \frac{b_p}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right]. \tag{5.33}$$

Our study will be divided in three parts: first, we inquire what can be learned from data that can already be found in the literature. Second, we introduce our own data and inspect some of its features. Third, we present the analysis of our own data and discuss results.

5.2.2 Exploratory study with PBC data

Unlike previously, we restrict ourselves to symmetric volumes N_S^4. The reason is that on asymmetric volumes $N_S^3 \times N_T$ we only obtained the total average plaquette (as a by-product of the Polyakov loop measurements) and did not distinguish between spatial and temporal contributions. The latter can be identified with chromoelectric and chromomagnetic contributions that, once calculated separately and then combined in a non-trivial way [220, 221], give access to thermodynamic quantities such as the energy density in QCD, which in turn is linked via Eq. (3.84) to the quantity we are eventually after, the gluon condensate.[46] In short, for symmetric volumes $N_S = N_T$ the situation is less complicated and one can

[46] For QCD at finite temperature, $N_T < N_S$, one can then define a temperature-dependent gluon condensate [222].

extract information from the total average plaquette alone. In the following, we will again use the shorthand notation $\{N_S; M\}$ to designate N_S^4 simulations up to M loops.

Our previous analysis suggested that the renormalon behavior of the plaquette in the lattice scheme should not be expected to set in before $n \sim 36$. As we have seen in Sec. 3.9.2, the data found in the literature falls short of this requirement. Nonetheless, it may be of interest for us as we also verified that the renormalon in the $d = 1$ static energy goes hand in hand with finite size effects $\sim N_S^{-1}$.

On these grounds, we expect for the plaquette finite size effects $\sim N_S^{-4}$,

$$p_n\left(N_S\right) = p_n - \frac{f_n(N_S)}{N_S^4} + \mathcal{O}\left(\frac{1}{N_S^6}, \text{lattice artefacts}\right). \tag{5.34}$$

Here, $p_n \equiv p_n\left(N_S = \infty\right)$ denotes the infinite volume result (we quoted the first three coefficients already in Sec. 3.9.2), and the coefficients f_n embody again the finite size effects at a typical scale of $1/(aN_S)$, with an entanglement of different orders due to the running of α up to the scale a^{-1} (cf. Sec. 5.1). In the sense that they are linked to the scale a^{-1} and no running is involved, lattice artefacts are a separate issue. With the Wilson gauge action being $\mathcal{O}\left(a^2\right)$ exact, we expect their contribution to be

$$\frac{h_n}{N_S^2} + \mathcal{O}\left(\frac{1}{N_S^3}\right). \tag{5.35}$$

To verify that Eq. (5.34) indeed captures the correct behavior of $p_n\left(N_S\right)$ at one and two loops, let us consult DLPT data up to $N_S = 64$, with an accuracy of at least 9 significant digits. PBC and Wilson gauge action are employed. The outcome is shown in Fig. 5.13, where we plot p_0, p_1 as functions of N_S^{-4}.

A closer look reveals that at LO, $f_0^{\mathrm{PBC}} = p_0$ and no further terms are needed.[47] At NLO, the fit formula

$$p_1^{\mathrm{PBC}}\left(N_S\right) = p_1 - \frac{f_1^{\mathrm{PBC}} + 2\frac{\beta_0}{4\pi} f_0^{\mathrm{PBC}} \ln\left(N_S\right)}{N_S^4} - \frac{h_1^{\mathrm{PBC}}}{N_S^2}, \tag{5.36}$$

describes the data very well. It appears as if the artefact coefficient h_1^{PBC} is arbitrarily small. We suspect it to be exactly zero at this order. Having convinced ourselves that Eqs. (5.34) and (5.35) capture the essential contributions to correctly parametrize the first two loops, let us turn to higher orders.

The NSPT data that can be found in the literature is also of PBC type. Recently [135], results for 20 loops and $N_S = 4, 8, 12, 16$ were published. They can be augmented by two older PBC simulations that were also computed with Wilson action but less perturbative loops, namely $\{24; 10\}$ from [133] and $\{32; 3\}$ from [196]. We denote the combined literature dataset by "DATAPBC".

[47] The superscript "PBC" shall indicate that FSE or artefact coefficients are not genuinely applicable, whereas the constant pieces p_n also hold for TBC.

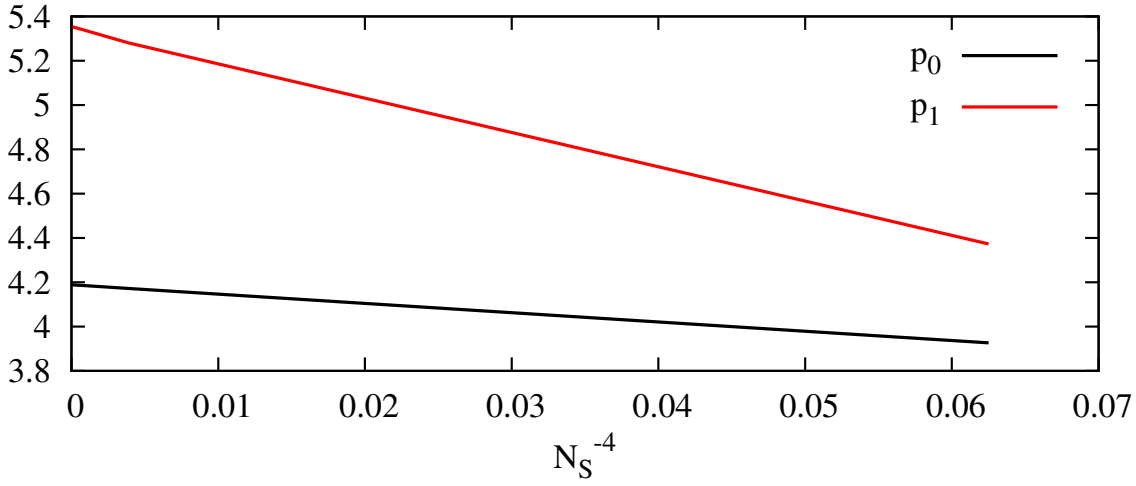

Fig. 5.13: The one- and two-loop coefficients p_0, p_1 of the plaquette as a function of N_S^{-4}. The numbers are DLPT results, where at least 9 significant digits may be assumed. The straight lines indicate that Eq. (5.34) is an adequate ansatz, to these orders.

d $\left(\text{ as in } N_S^{-d}\right)$	χ^2_{red} (run)	χ^2_{red} (no run)
1	144.08	73.82
2	75.23	23.83
3	26.41	6.67
4	4.85	8.64
5	4.19	17.52
6	12.88	26.88
7	23.11	34.53

Table 5.2: Exploratory fits to DATAPBC and the resulting χ^2_{red} as a measure of the goodness of fit. All fits have two parameters per order n. The second column lists χ^2_{red} for generalizations of Eq. (5.34) for arbitrary dimension d, the third column for similar generalizations of Eq. (5.35).

Let us make a survey of exploratory fits to DATAPBC without (theoretical) prejudice: we try out various parametrizations of FSE with leading term $\propto N_S^{-d}$, only one of which is Eq. (5.34) with $d = 4$. Per order n, we thus have two parameters, the constant piece p_n plus the FSE term. The resulting χ^2_{red} (as a measure of the goodness of fit) are shown in the second column of Table 5.2. The numbers indicate that the parametrization works best for $d \sim 4-5$, both higher and, most notably, lower values of d are clearly ruled out by the data. The third column lists χ^2_{red} for fits in which we also have two parameters per order n, but, instead of FSE terms involving the running coupling, we use terms generalizing Eq. (5.35) where no running is involved. In this case, χ^2_{red} is minimal at $d = 3$ but clearly above the FSE terms. It shows that the running of the coupling is indispensable also in the plaquette case. With only 4 DATAPBC points per order for $n > 9$ and zero points beyond $n = 19$, we refrain from fine-tuning the survey and turn to our own data.

loop order	N_S
$\mathcal{O}\left(\alpha^5\right)$	$11, 13$
$\mathcal{O}\left(\alpha^{20}\right)$	14
$\mathcal{O}\left(\alpha^{30}\right)$	$12, 40$
$\mathcal{O}\left(\alpha^{35}\right)$	$3, 4^{(\mathbf{5})}, 5, 6^{(\mathbf{12})}, 9, 10^{(\mathbf{12})}, 28^{(\mathbf{35})}, 30$
$\mathcal{O}\left(\alpha^{40}\right)$	$7, 8^{(\mathbf{12})}, 16^{(\mathbf{12})}, 18, 20, 22, 24, 32$

Table 5.3: The first column states to which order in α the coefficients of $p_n(N_S)$ were computed for each specific symmetric lattice volume N_S^4. Altogether, we have considered 21 different lattice volumes. Bracketed superscripts label those lattices for which we performed the extrapolation $\epsilon \to 0$ and indicate the loop number up to which we did so. Otherwise, simulations at a single $\epsilon = 0.05$ are implied.

5.2.3 The influence of the stochastic time step

In Table 5.3, we list the various values of N_S we used as well as the perturbative order up to which we measured. Throughout, we have employed TBCxyz and the Wilson gauge action. As far as the loops are concerned, our simulations range from $\mathcal{O}\left(\alpha^5\right)$ to $\mathcal{O}\left(\alpha^{40}\right)$. The fact that we did not calculate all geometries up to the same perturbative order is due to the restrictions of our computational capacities and reflects the synergies with the Polyakov loop study. Orders as high as $\mathcal{O}\left(\alpha^{40}\right)$ are not only at the limit of present-day computing powers, they also strain the numerical stability of both NSPT simulations and fitting procedures. To be safe, we only include orders as high as $\mathcal{O}\left(\alpha^{35}\right)$ in our analysis.

As before, most of our simulations were carried out at a single stochastic time step $\epsilon = 0.05$, constituting a data set again denoted as "DATA50". To estimate the effect of varying the step size, we simulated a fairly large lattice volume $\{28; 35\}$ at three values of $\epsilon = 0.04, 0.05, 0.06$. In Fig. 5.14 we give an overview of the plaquette's dependence on ϵ by depicting all odd loop numbers. The simulation results for the three time steps (blue) are shown together with the purely quadratic fit $y = \text{const.} + b\epsilon^2$ and fit result (red). For convenience, we normalize the axes as follows: the horizontal axis yields ϵ^2 in units of 10^{-3}, while the vertical axis is normalized to the corresponding $p_n\left(\epsilon = 0.05\right)$. We conclude that at intermediate orders $n \approx 23$ the ϵ dependence is more pronounced than at low and very high orders. More precisely, at low orders, our data are very accurate and indicate a particularly flat extrapolation in this regime.[48] At very high orders, however, the situation is less conclusive due to the lack of accuracy. Overall, the ϵ dependence is mild: as the figures' vertical axes indicate, the normalized relative error between the central values of $p_n\left(\epsilon = 0.05\right)$ the fitted value $p_n\left(\epsilon = 0\right)$,

$$d_\epsilon(n) = \left| \frac{p_n\left(\epsilon = 0.05\right) - p_n\left(\epsilon = 0\right)}{p_n\left(\epsilon = 0.05\right)} \right| , \tag{5.37}$$

never exceeds 10 percent.

[48] One noteable exception is $\mathcal{O}\left(\alpha^2\right)$, which is not depicted here.

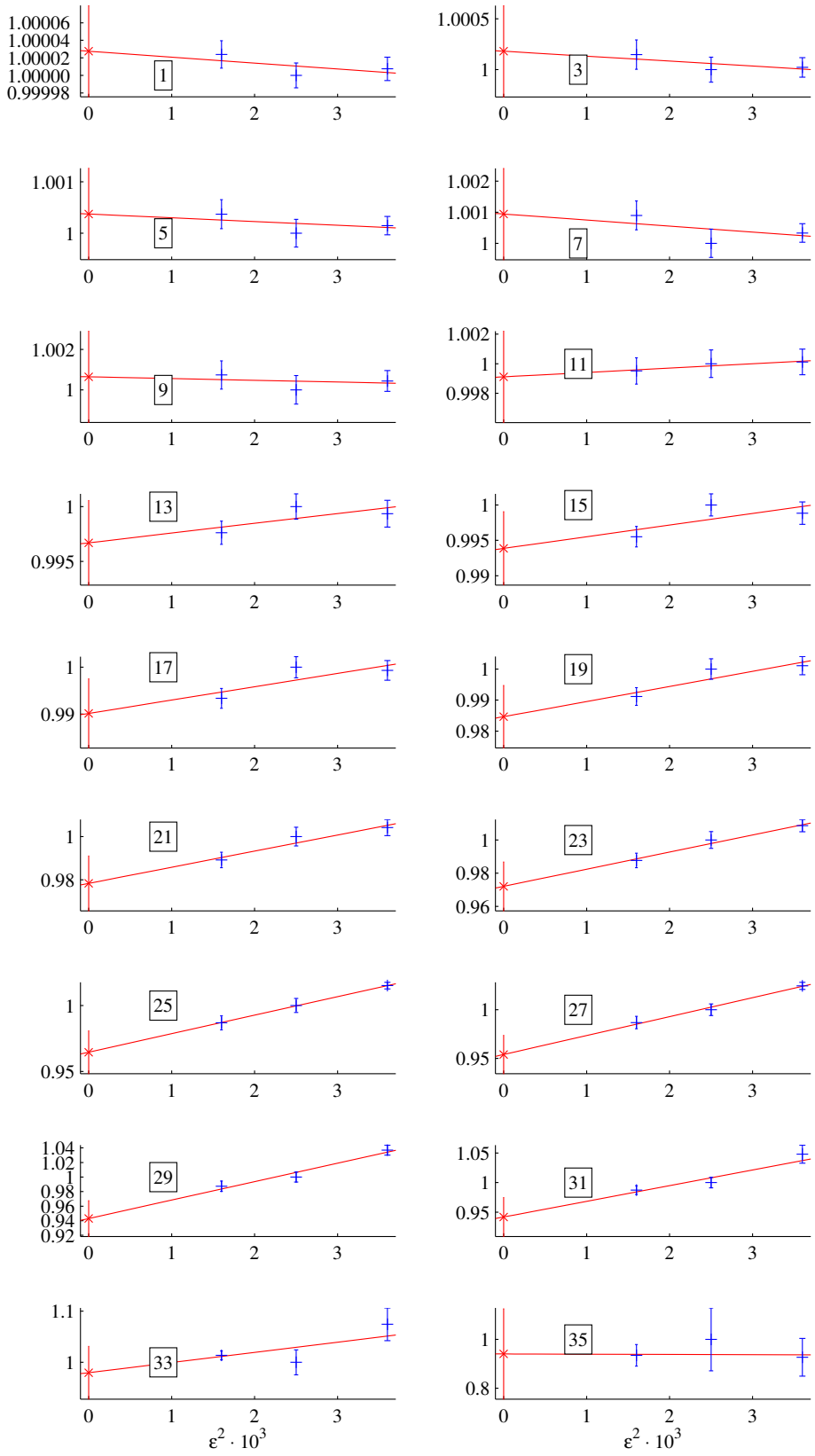

Fig. 5.14: Overview of ϵ effects in plaquette simulations on a 28^4 lattice. Numbers in boxes refer to the associated loops. Details are given in the text.

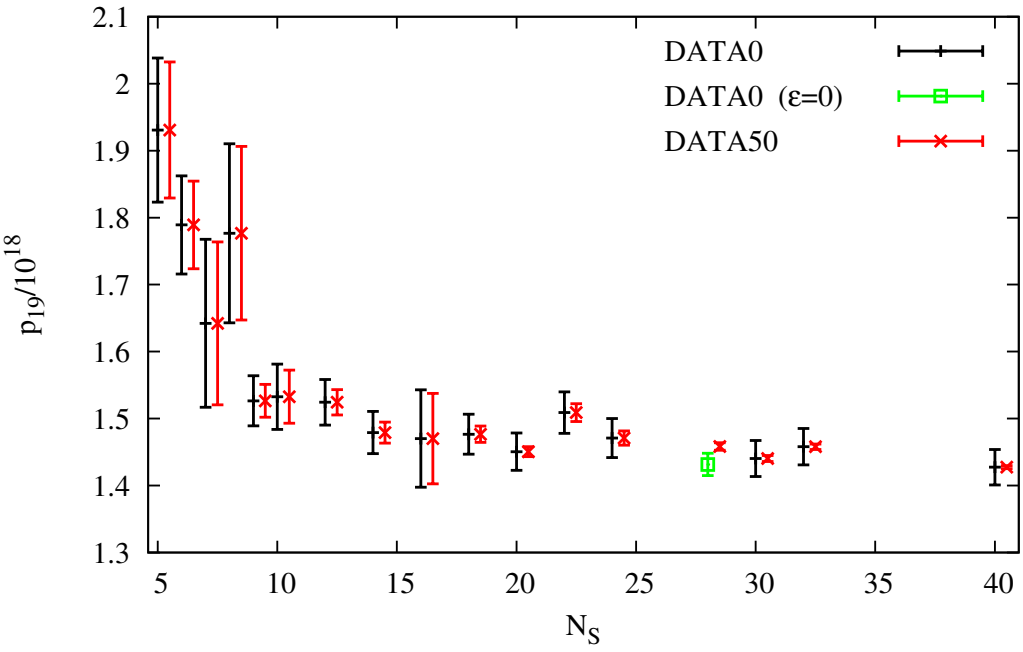

Fig. 5.15: DATA0 (left, black) versus DATA50 (shifted 0.5 units to the right, red) for the plaquette at 20 loops, p_{19}. Observe the notable difference between the two datasets on big lattice volumes: here, the DATA50 error (the statistical uncertainty on $p_{19}(\epsilon = 0.05)$) is very small, so the inclusion of d_ϵ [cf. Eq. (5.37)] leads to drastic increase. On small volumes, however, the dominant source of error is the DATA50 error itself, hence DATA0 differs only marginally. Note the exception at $N_S = 28$, where DATA0 yields the true (extrapolated) $\epsilon = 0$ result (shown in green).

It stands to reason that we repeat the procedure of Sec. 5.1 and approximate the uncertainty due to the omitted ϵ extrapolation for a given simulation by adding the d_ϵ obtained from $\{28; 35\}$ in quadrature to the statistical errors on p_n ($\epsilon = 0.05$). In the few cases where the ϵ extrapolation has been carried out, we instead use the extrapolated p_n ($\epsilon = 0$) result and the associated error. The resulting, second data set is being referred to as "DATA0".

Let us examine the two data sets: in Fig. 5.15, we contrast for the 20-loop plaquette coefficient, p_{19}, DATA0 values (left, black) with DATA50 values (right, red). We see that DATA50 is extremely accurate on big lattices. The inclusion of d_ϵ inflates the error and, consequently, DATA0 exhibit larger uncertainties. This effect is desirable, as Fig. 5.14 already illustrated: on volumes as large as $N_S = 28$, the statistical error on results $p_n(\epsilon)$ at different ϵ is small enough to see the dependence on the step size. It is a conservative approach to use the same d_ϵ obtained from $\{28; 35\}$ also for small volumes, since they come with statistical errors that are too large to resolve ϵ effects. In any case, at small N_S statistical errors are the dominant source of error and the quadrature ensures that it is a minor change from DATA50 to DATA0 in this regime.

DATA0 at $N_S = 28$ is shown in green in Fig. 5.15 to indicate that here it yields the true (extrapolated) $\epsilon = 0$ result. Its good agreement with DATA50 within errors underlines once more the flatness of the second-order integrator, and it fits in nicely with the other DATA0 points (black).

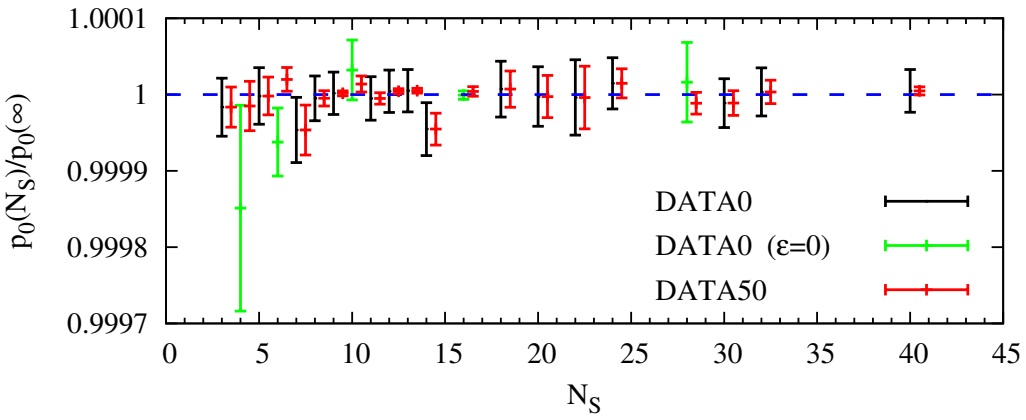

Fig. 5.16: DATA0 (left, black) versus DATA50 (shifted 0.5 units to the right, red) for the plaquette at LO, p_0. DATA0 is shown in green instead of black where we simulated at several time steps ϵ such that DATA0 yields the true (extrapolated) $\epsilon = 0$ result. The dashed blue horizontal line corresponds to the infinite-volume result $p_0(\infty)$ from DLPT, which also serves to normalize the vertical axis. Note that all DATA50 values agree with DLPT within 2 σ, and a few very accurate DATA50 values such as $N_S = 13$ all comply with DLPT well within $\mathcal{O}\left(10^{-5}\right)$ precision.

For a more quantitative analysis at one and two loops, see Figs. 5.16 and 5.17 and the associated captions. The plots include several DATA0 values with (extrapolated) $\epsilon = 0$ results as well as the accurate DLPT value at infinite volume. In summary, our account of the error due to ϵ indeed seems plausible.

5.2.4 Fits

One consequence of using TBCxyz instead of PBC is that the one-loop behavior is flat, $p_0^{\mathrm{TBC}}(N_S) = p(\infty) \equiv p_0$. This is evident from one-loop DLPT data, and we may use them as a reference to put our own data to a precision test. In Fig. 5.16, we contrast our $p_1(N_S)$ values with the infinite-volume value from DLPT. We see agreement within errors down to the smallest lattice volumes, so FSE or artefact terms are truly absent at LO.

One would assume that already at NLO these terms should start to show up, and indeed this is the case (cf. Fig. 5.17). The fall-off at low N_S together with the plateau at large N_S points towards a dominant N_S^{-4} dependence. We saw that in PBC DLPT data up to NLO the N_S^{-2} artefact term indeed vanishes or is very small. This pattern need not extend to three loops and beyond, where a N_S^{-2} artefact term may very well contribute. Note furthermore that the Wilson gauge action permits artefacts in powers of N_S^{-2}. We will come back to this point shortly. Fig. 5.17 also reveals that the step size errors are only weakly volume-dependent, justifying the way we obtain DATA0, at least to low orders.

As a consequence of the aspects discussed so far, we devise the following fit strategy to begin with (and will add to it below):

1. We attempt a global (simultaneous) fit to our data with two independent parameters, N_s and n.

2. Since the LO is a constant, we exclude it from the fit and consider 2–35 loops.

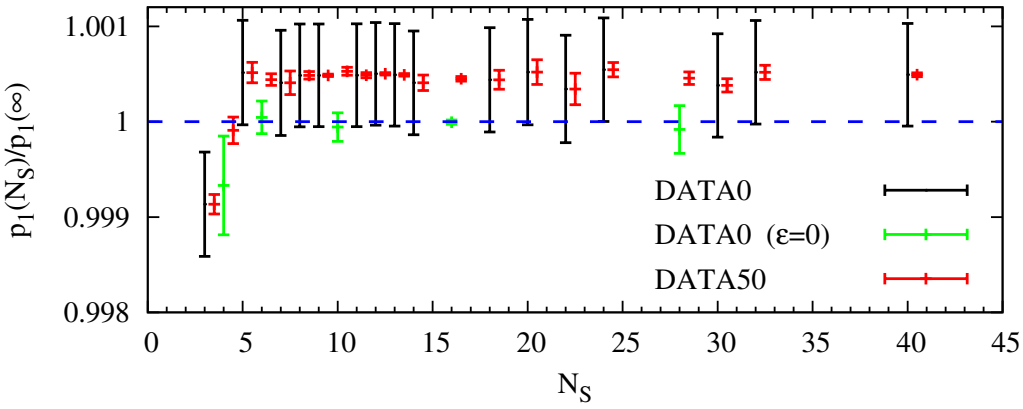

Fig. 5.17: DATA0 (left, black) versus DATA50 (shifted 0.5 units to the right, red) for the plaquette at NLO, p_1. DATA0 is shown in green instead of black where we simulated at several time steps ϵ such that DATA0 yields the true (extrapolated) $\epsilon = 0$ result. The dashed blue horizontal line corresponds to the infinite-volume result $p_1(\infty)$ from DLPT, which serves to normalize the vertical axis. Note the shift from DATA50 to DATA0 which is similar in sign and magnitude for those values where the $\epsilon = 0$ result is known. It indicates that the dependence on ϵ does not change significantly with N_S. In hindsight, this justifies the use of d_ϵ obtained on $\{28; 35\}$ for all volumes. A further confirmation is the agreement of black DATA0 values with DLPT in the plateau region ($N_S \gtrapprox 5$).

3. For those orders where the infinite-volume diagrammatic result is known [Eq. (3.77)], we fix the constant piece accordingly to isolate the FSE and artefact contributions.[49]

For a start, we implement only FSE besides the constant piece, which yields a goodness of $\chi^2_{\text{red}} \approx 2.36$ and definitely needs improvement. To rule out unwanted higher-order effects, which are expected to have their biggest impact at very small N_S,

4. we exclude our smallest lattice volume, $\{3; 35\}$, from the fit.

This leads to $\chi^2_{\text{red}} \approx 1.70$ (referred to as "**Fit 1**" below), and it is reasonable to adopt this restriction for good. A further elimination of data points, however, reduces χ^2_{red} only slightly ($\chi^2_{\text{red}} \approx 1.63$ for $N_S > 4$). To achieve a good fit, we must modify our fit function. Beyond NLO, we may admit the leading artefact term $\propto N_S^{-2}$ and think that it cannot be avoided (having convinced ourselves that a global fit with the only artefact term being $\propto N_S^{-4}$ makes little difference in the fit). Apart from that, there can be subleading (artefact or FSE) terms $\propto N_S^{-6}$ in principle, but they are heavily suppressed since the bulk (and most precise) part of our data has $N_S \geq 10$. What is more, a fit including such terms (**Fit 2**) fails to describe our data at large orders, $n \approx 30$. Hence,

5. we finally introduce the leading artefact term $\propto N_S^{-2}$ for three loops and beyond. In total, we then have three fit parameters per order.

In this last step, we obtain a satisfactory $\chi^2_{\text{red}} \approx 1.30$ (**Fit 3**).

[49] To be able to do so, we convinced ourselves with order-by-order fits (in which the resulting FSE fit parameters are propagated to the fit of following order) that we are able to reproduce the infinite-volume results.

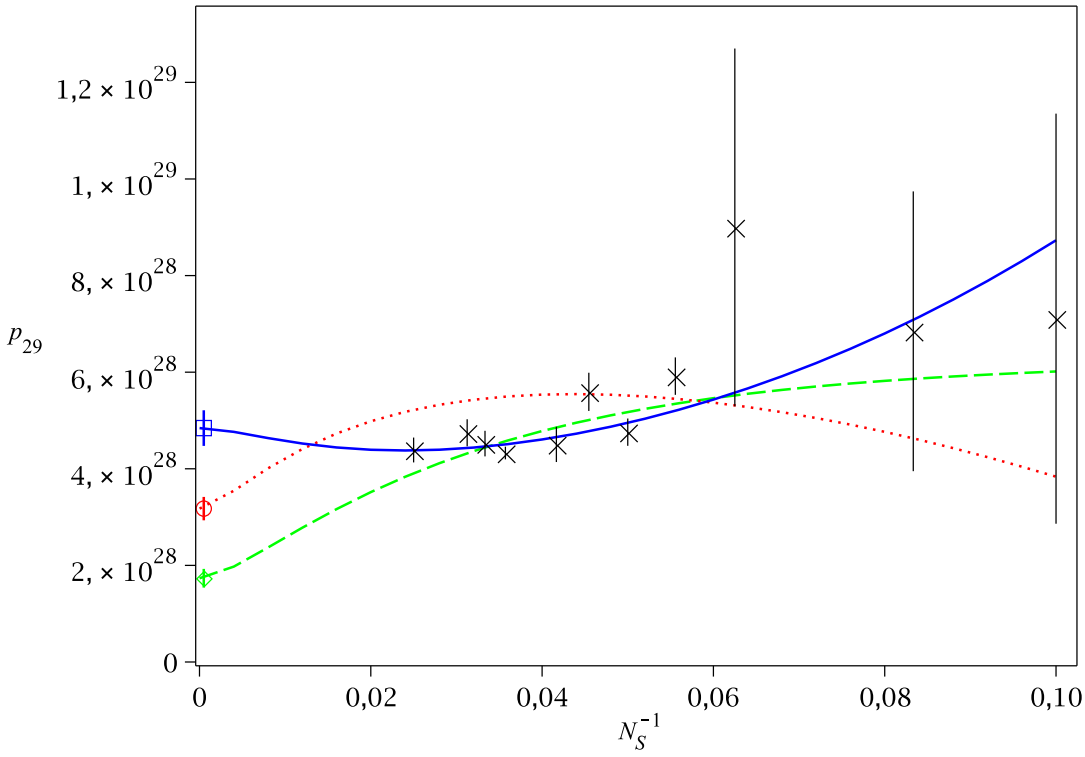

Fig. 5.18: Our data versus Fit 1 (dashed, green), Fit 2 (red, dotted), and Fit 3 (solid, blue) for the plaquette at 30 loops, $p_{29}(N_S)$. For illustration, we show the resulting infinite-volume coefficients and associated errors.

At large N_S, our data forms a narrow plateau up to very large perturbative orders, which opens a second way to discriminate between parametrizations. We exemplarily compare in Fig. 5.18 the 30-loop coefficients $p_{29}(N_S)$ that follow from the best-fit parameters of Fits 1, 2 and 3 with our data. Fit 3 both reproduces the plateau and delivers an (infinite-volume) p_{29} in its vicinity, whereas all previous attempts fail to do so. We conclude that Fit 3 is our final result and the following discussion will be based upon it.

Let us first see whether we detect the notorious plaquette renormalon, which has not been confirmed so far despite the longstanding efforts of other groups (cf. Sec. 3.9.2). We adopt the definition of Eq. (5.33) and show our ratio curve along with the LO, NLO and NNLO expectation in Fig. 5.19. We have clear evidence for the onset of factorial growth in our infinite-volume coefficients p_n around $n \approx 27$, in agreement with the (conjectured) plaquette IR renormalon of dimension $d = 4$. More quantitatively, ratios 25 - 34 all agree with the LO expectation within (at most) 1.35 σ. The last value we compute, ratio 34 comes with a relative error of about 10 percent, while preceeding orders lie significantly below, so we emphasize the accuracy with which we can compute ratios up to highest orders. In doing so, we have exploited the covariance matrix of our fit parameters. When computing ratios of subsequent plaquette coefficients, correlations are such that their inclusion reduces the error. Even without this step, the plaquette renormalon would be seen unambiguously with a precision of $\mathcal{O}(15\,\%)$ in the relevant regime. We recall that only the first three

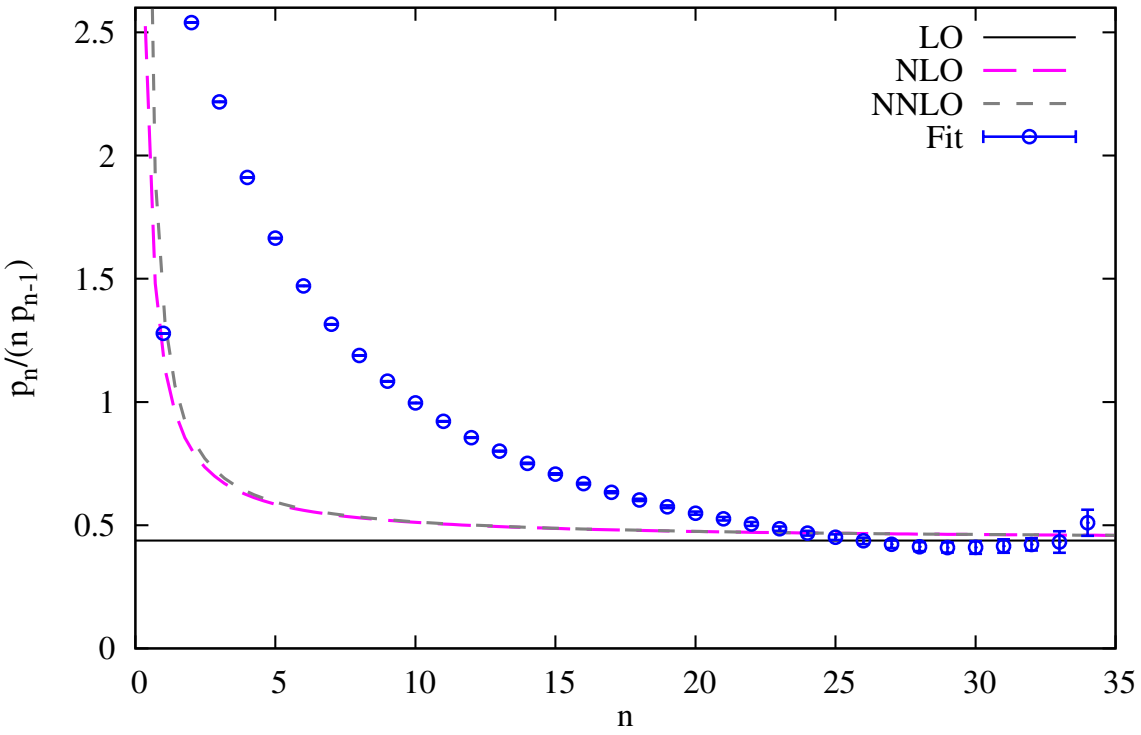

Fig. 5.19: Ratios of plaquette coefficients converging to a $d = 4$ IR renormalon as predicted. For a close-up view of the relevant region of large orders n, see also Fig. 5.20.

coefficients of the β function are known in the lattice scheme. For an exact treatment of the running coupling in our fit, we would need to know the coefficients beyond 30 loops. We revert the argumentation and see whether our results are stable under exclusion of β_i which we already know. Let us therefore repeat our fit twice, with $\beta_2 = 0$ and $\beta_1 = \beta_2 = 0$. The goodness of fit practically does not change, as previously $\chi^2_{\text{red}} \approx 1.30$ in both cases. The outcome for the ratios is shown in Fig. 5.20. We conclude that the sign of renormalon growth is stable under exclusion of higher β_i, in analogy to our findings in Sec. 5.1 for the static energy.

To complete our ratio analysis, we turn to the FSE and artefact coefficients, f_n and h_n, to see if and what pattern they follow. Although one needs to incorporate them to obtain a good fit, the coefficients of the $\mathcal{O}\left(N_S^{-4}\right)$ FSE contribution turn out to be weakly determined and so are the ratios, as Fig. 5.21 illustrates.[50] At large orders n, the FSE ratio curve does not yield the $d = 4$ renormalon pattern, in contrast to the behavior we observed for the $d = 1$ static energy. This discrepancy may be resolved as follows: for the static energy, the $\mathcal{O}\left(N_S^{-1}\right)$ FSE represented the leading correction in N_S. The situation is more complicated in the plaquette case, where the $\mathcal{O}\left(N_S^{-2}\right)$ artefact term is the leading correction and dominant over the FSE in the sense that they are only $\propto N_S^{-4}$. Still, FSE include the running coupling with large logarithm terms at large orders, while the artefact term does not. Therefore, it is not obvious at all if either one of them should prevail.

[50] See also Appendix A.5, where the values are tabulated.

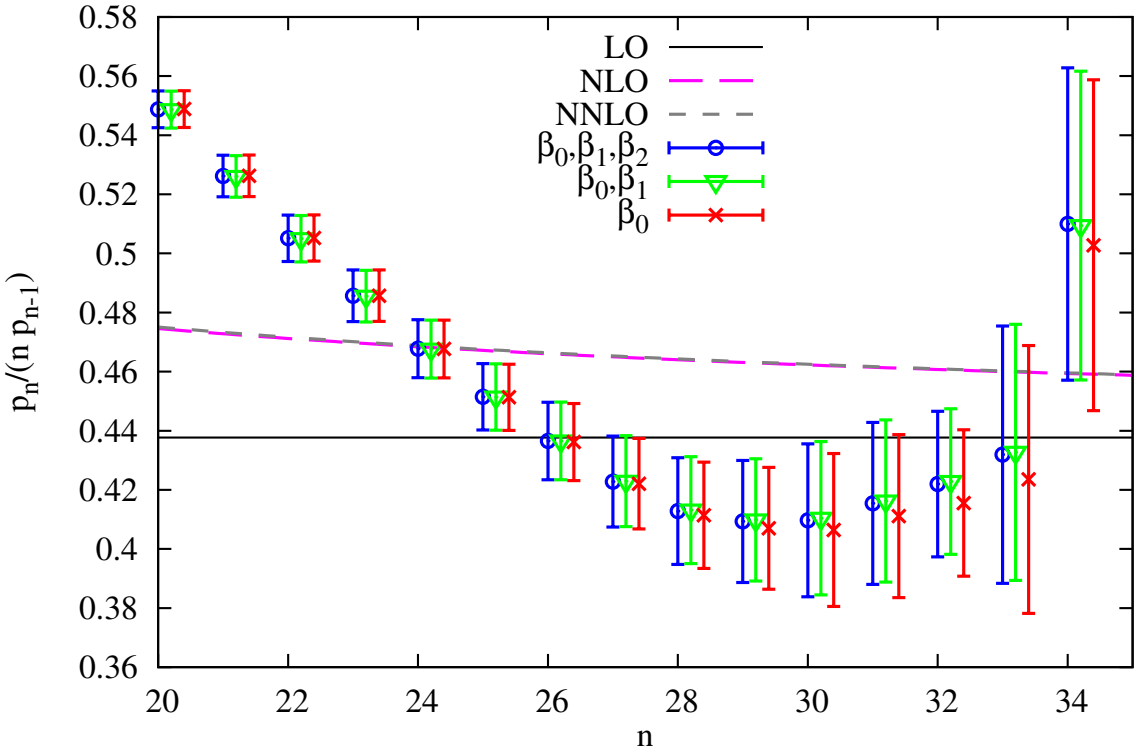

Fig. 5.20: Comparison of ratio curves for fits with $\beta_0, \beta_1, \beta_2$ (blue), β_0, β_1 (green) and β_0 only (red).

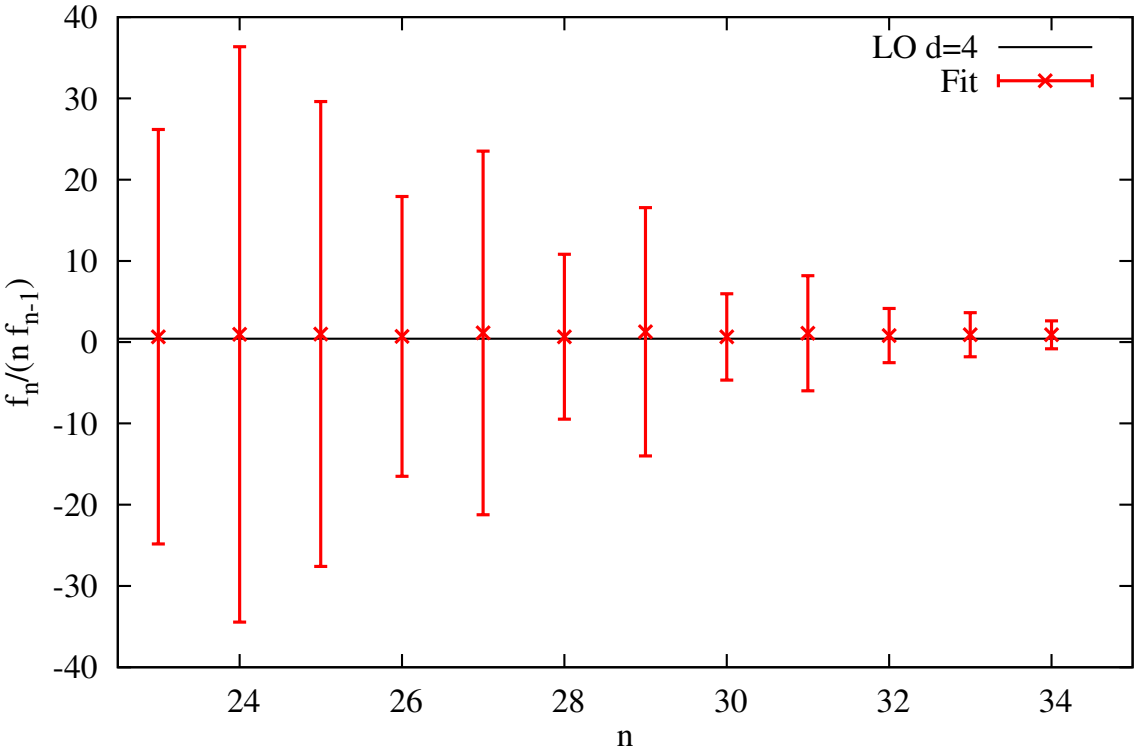

Fig. 5.21: Ratios of FSE coefficients versus the theoretical LO $d = 4$ IR renormalon prediction.

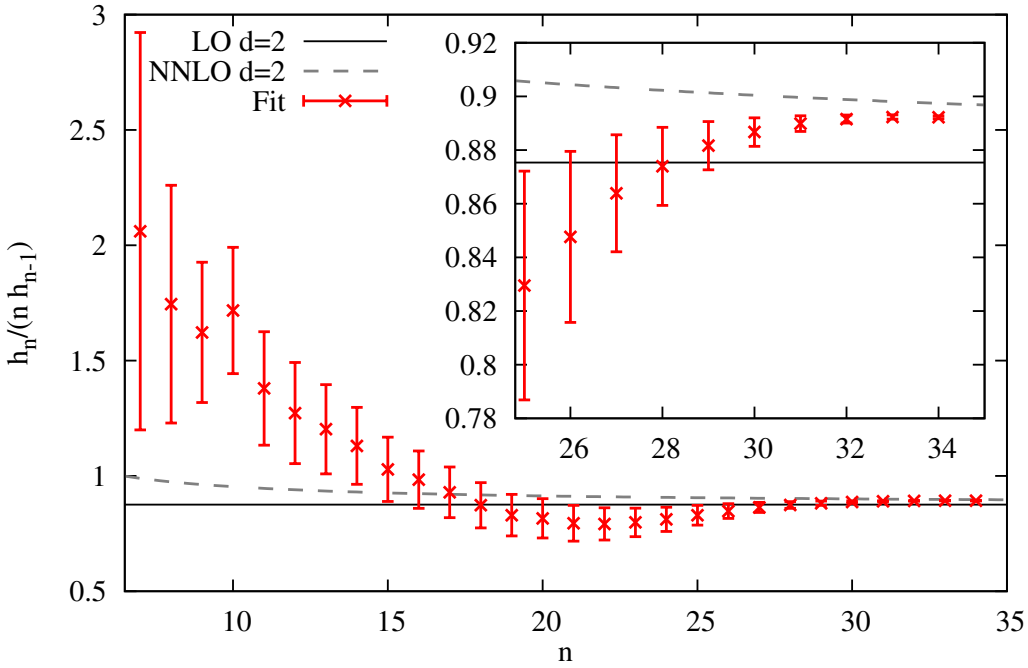

Fig. 5.22: Ratios of artefact coefficients, converging to a $d = 2$ IR renormalon.

Surprisingly, the artefact coefficients h_n exhibit a $d = 2$ IR renormalon, as Fig. 5.22 demonstrates. We discussed at length in Chap. 3 that a $d = 2$ IR renormalon is forbidden in the continuum. Nonetheless, things are more opaque on the lattice and our findings may partly explain the recurrent speculations on a $d = 2$ plaquette renormalon in previous NSPT studies (cf. Sec. 3.9.2 as well as Sec. 6.2 of [37] and the references therein). An (otherwise identical) NSPT study with a Symanzik-improved gauge action would probably resolve the puzzle because we suspect that the impact of the artefact term should be heavily reduced in that case. We repeat: our main result is that the plaquette coefficients p_n comply with the conjectured $d = 4$ plaquette renormalon. In addition, we find a $d = 2$ IR renormalon-like behavior in the subleading correction.

As mentioned, the computation of ratios according to Eq. (5.33) is favored by the correlation between coefficients of subsequent orders. Also, any dependence on the plaquette normalization N_p cancels. It is interesting to extract N_p, though, because after the onset of renormalon growth the plaquette coefficients are entirely dictated by Eq. (5.32). We show $N_p\,(n)$ in Fig. 5.23 as well as the outcome of the aforementioned fits omitting β_2 (and β_1), which also requires to use Eq. (5.32) up to the corresponding order.

Naturally, the errors on N_p reflect the uncertainty in the coefficients p_n themselves. It seems reasonable to account for our ignorance of higher β_i by measuring the relative difference between the central values of two fits: we take the difference between our main fit and the fit with β_2 set to zero. We then add this difference in quadrature to the fit-induced relative errors, and arrive at a total uncertainty for our coefficients p_n, f_n, h_n. All coefficients and the corresponding ratios can be found in Appendix A.5.

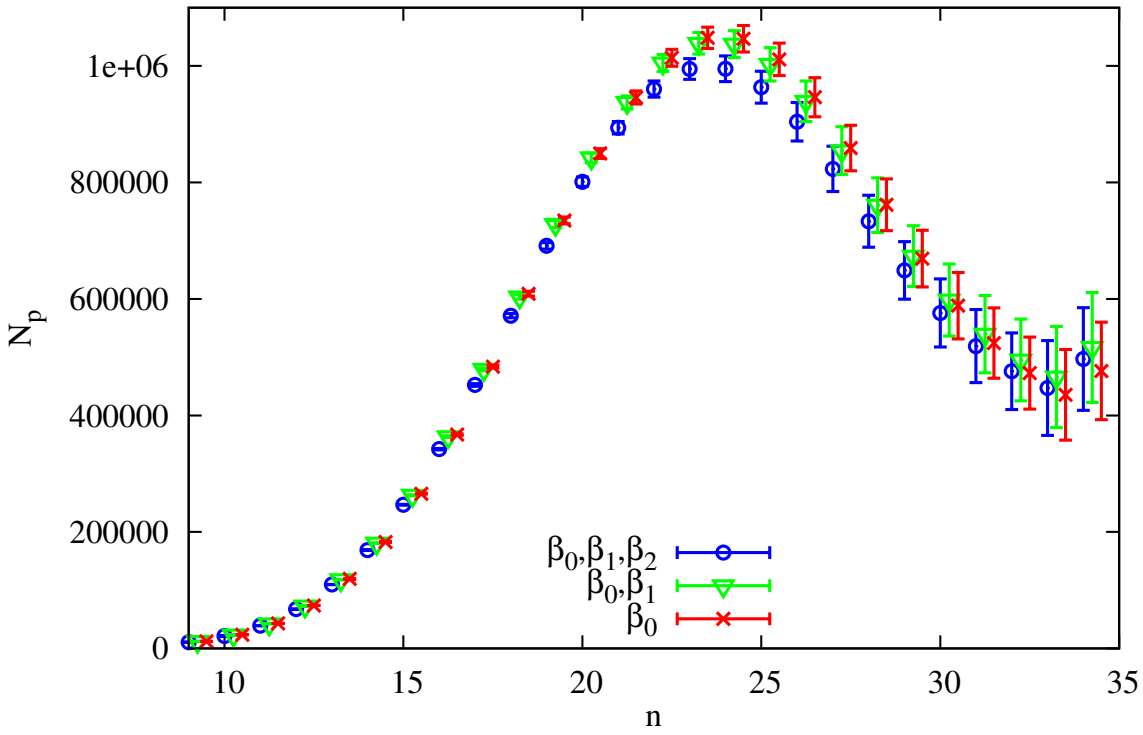

Fig. 5.23: Comparison of the plaqette normalization N_p for fits with $\beta_0, \beta_1, \beta_2$ (blue), β_0, β_1 (green) and β_0 only (red).

As for N_p itself, the plaquette situation is less comfortable than in Sec. 5.1, where we went well beyond the renormalon onset and obtained a clear plateau for N_m. Here, we just about reach the renormalon onset, so whatever value of N_p we establish should be treated with due caution. We settle for an average of $N_p(n)$ from $n = 30$ to $n = 34$. As an uncertainty, we quote the largest relative error in this range and arrive at

$$\begin{aligned} N_p^{\mathrm{L}} &= 4.85(88) \times 10^5 \,, \\ N_p^{\overline{\mathrm{MS}}} &= 0.704(13) \,. \end{aligned} \tag{5.38}$$

Note that the normalization in the $\overline{\mathrm{MS}}$ scheme is $\mathcal{O}(1)$, as it was the case for the static self-energy [Eq. (5.29)]. The result for the static self-energy normalization is much more solid, however, and the reasons were discussed above. The scheme-independent ambiguity can be computed in analogy to Eq. (5.28), it amounts to

$$N_p^{\overline{\mathrm{MS}}} \left(\Lambda_{n_f=0}^{\overline{\mathrm{MS}}} \right)^4 = 0.00201(37)(66)\ \mathrm{GeV}^4 \,. \tag{5.39}$$

Finally, we use our plaquette coefficients p_n up to unprecedented 35 loops together with non-perturbative plaquette results to distill the gluon condensate using Eq. (3.75). The non-perturbative data is taken from [223]. Besides its statistical error, we have to consider a volume-related uncertainty because of the finite lattice extents used in that work, $N_S = 16$ and $N_S = 32$. As Fig. 5.24 illustrates for different lattice β, the deviations are small

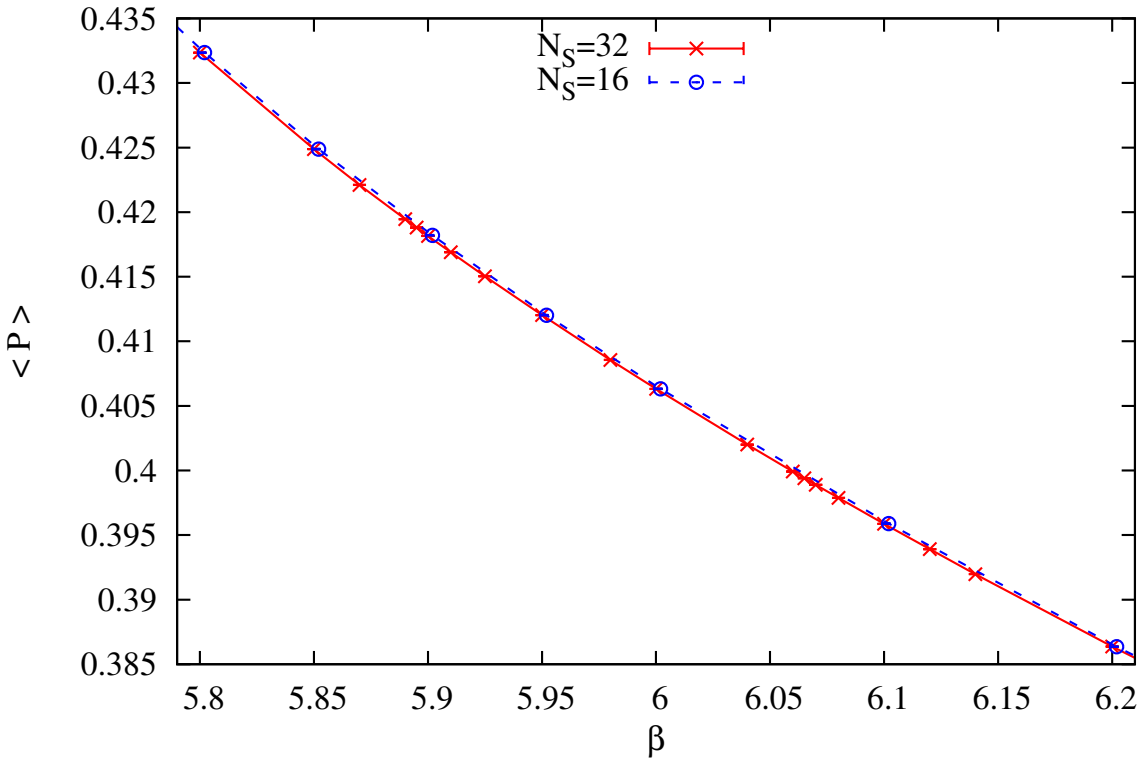

Fig. 5.24: Comparison of the non-perturbative plaquette $\langle P \rangle$ for $N_S = 32$ (red, solid) and $N_S = 16$ (blue, dashed) from [223] for different values of lattice β.

and we will see later that the other ingredients of our analysis come with much larger errors. Most importantly, we will need to translate the values of lattice β used in [223] into the corresponding lattice spacings a. Instead of Eq. (A.9), it is better to use the phenomenological conversion given in [219],

$$a = r_0 \exp\left[-1.6804 - 1.7331(\beta - 6) + 0.7849(\beta - 6)^2 - 0.4428(\beta - 6)^3\right] , \qquad (5.40)$$

with Sommer scale $r_0 = 0.5\,\text{fm}$. Nonetheless, its applicability is limited to a finite range around $\beta \approx 6$, a fact which we will keep in mind.

For the β values of [223], we can resum our perturbative plaquette series in powers of $\alpha = 3/(2\pi\beta)$. First, we show individual contributions $p_n \alpha^{n+1}$ in Fig. 5.25 for different values of α. It is no surprise that we encounter minimal terms close to the LO theoretical prediction, Eq. (3.36). It has to be this way for the $d = 4$ IR renormalon we encountered. Also consistent with theory is that the successive contributions start to rise again once $n > n_0$. We pointed out already that it makes no sense resumming the series beyond the minimal term.

For $\beta = 6$, the plaquette series resummed to various orders n is shown in Fig. 5.26, along with the non-perturbative $\langle P \rangle$. In correspondance to the above, we see how the resummed series starts growing faster for $n > n_0$. Eventually, it will diverge.

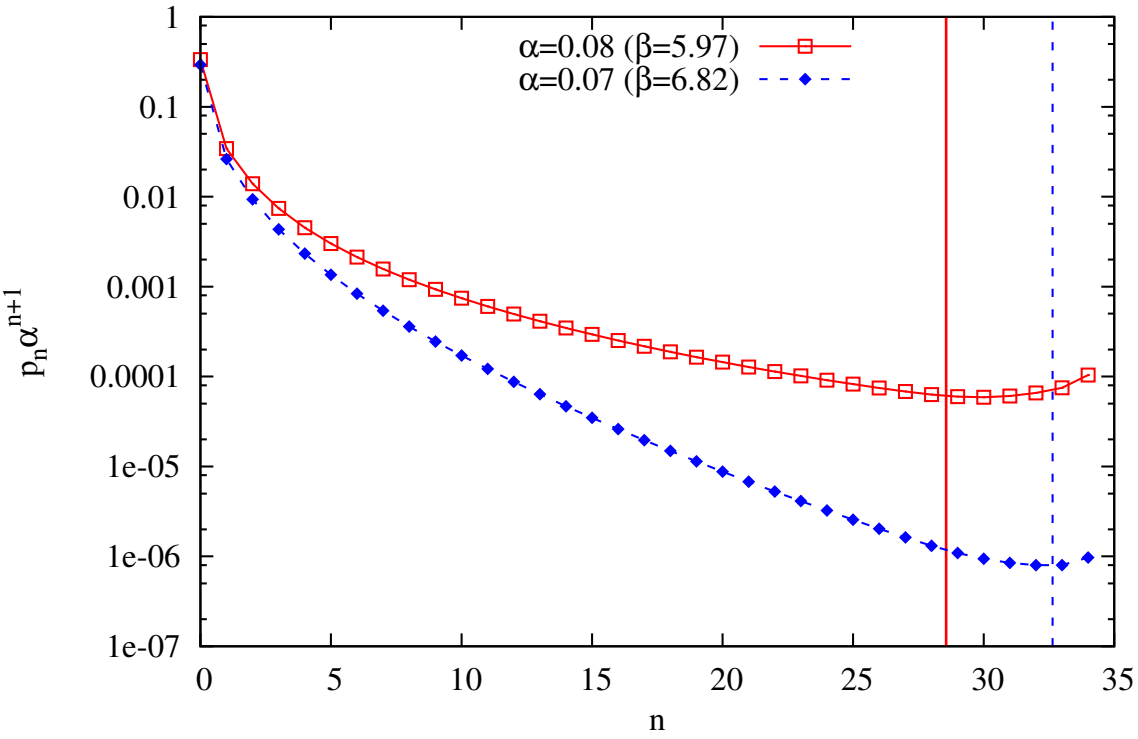

Fig. 5.25: Successive contributions $p_n\alpha^{n+1}$ to the resummed plaquette for $\alpha = 0.08$ (red, solid) and $\alpha = 0.07$ (blue, dashed). Vertical lines in alike colors indicate the LO theoretical expectation for the order n_0 at which the minimal term $p_{n_0}\alpha^{n_0+1}$ should occur.

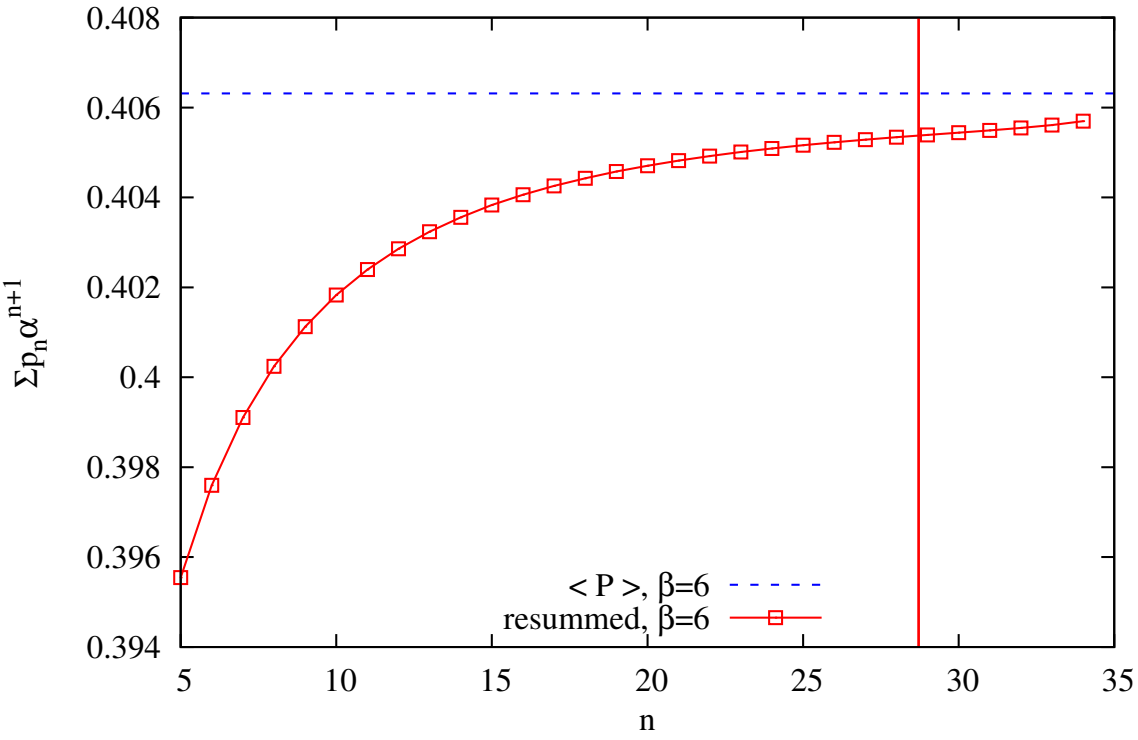

Fig. 5.26: Our plaquette series resummed up to order n versus the non-perturbative $\langle P \rangle$. The vertical line indicates the location of the LO theoretical n_0 at $\beta = 6$.

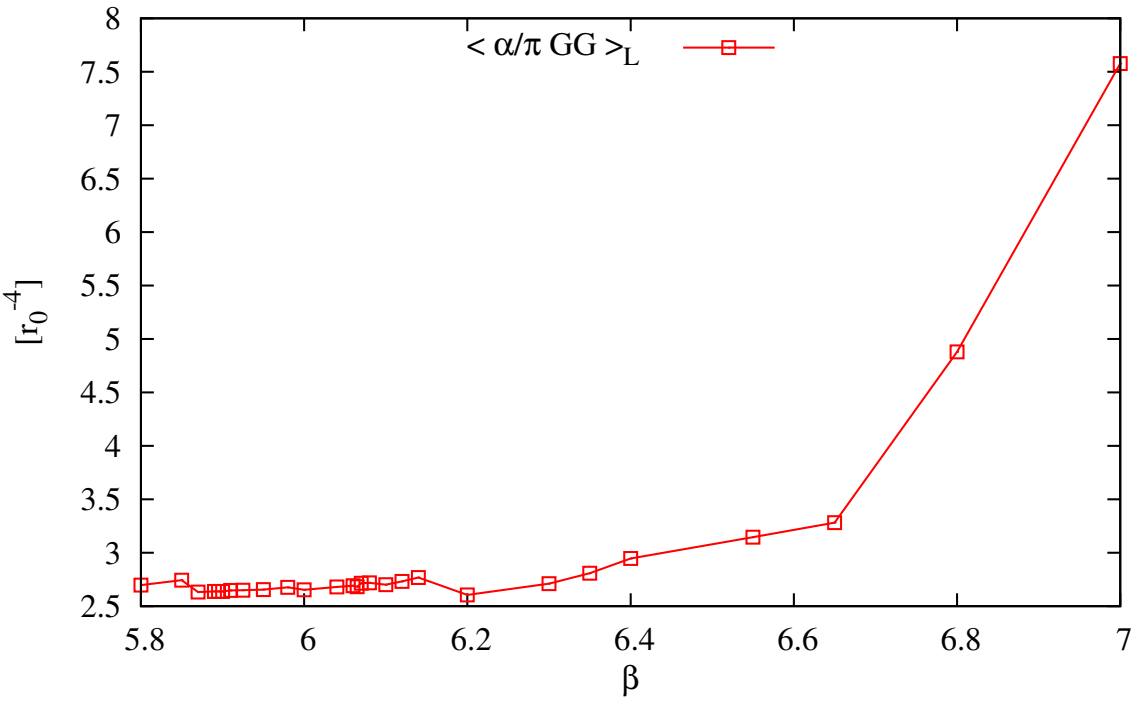

Fig. 5.27: The gluon condensate as a function of lattice β. For each value of β, the perturbative series has been resummed up to the (empirical) minimal term. Note the formation of a plateau at low β.

In a final step, we extract the gluon condensate on the lattice, $\langle \frac{\alpha}{\pi} G^2 \rangle_{\mathrm{L}}$, from the difference of non-perturbative $\langle P \rangle$ and resummed perturbative data using Eq. (3.75). To this end, we stop the resummation after the minimal term as it is found empirically for each distinct value of β. The result is shown in Fig. 5.27 for different values of β at which $\langle P(\beta) \rangle$ is computed in [223].

It is remarkable that the resulting curve becomes nearly flat at $5.8 \leq \beta \lessapprox 6.3$, indicating that the gluon condensate is independent of β in a large range of lattice spacings. In other words, $\langle \frac{\alpha}{\pi} G^2 \rangle_{\mathrm{L}} \approx \langle \frac{\alpha}{\pi} G^2 \rangle$ in this regime. The fact that large β do not fit into this interpretation has two possible explanations. First, Eq. (5.40) is only valid for $\beta \approx 6$, beyond that it entails large uncertainties. Second, the condensate can only be extracted if $\langle P(\beta) \rangle$ is obtained for β low enough that the non-perturbative simulation is in the confined phase. To rule out both unwanted effects, we restricted ourselves to a symmetric interval $\beta = 6 \pm 0.2$. We average the resulting values for the gluon condensate in this range. This will give our final number for the gluon condensate. We separately account for different sources of error: instead of central values (Fig. 5.27), we calculate the condensate using the upper 1 σ bounds of perturbation theory (PT) and $\langle P \rangle$ (MC) to estimate the associated uncertainties. The result is shown in Fig. 5.28, together with the corresponding size of the empirical minimal term $p_{n_0} \alpha^{n_0+1}(\beta)$. Our explanation why it is not as flat as the other curves is that we we pick n_0 among discrete n. It quantifies the ambiguity in the perturbative expansion itself and yields our third and so far biggest source of error. Still, it is more than a magnitude smaller than the condensate in the β-range under inspection.

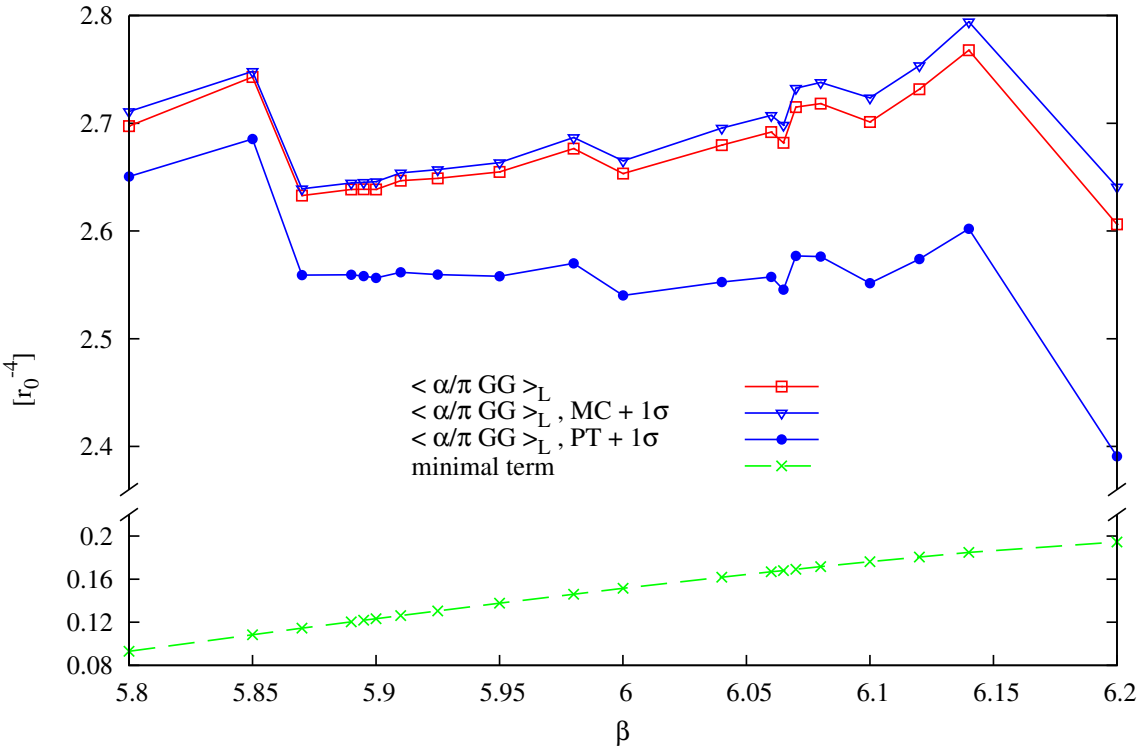

Fig. 5.28: The gluon condensate as a function of lattice β in the plateau region. The solid red line corresponds to a resummation up to the (empirical) minimal term in our data. As a variation, solid blue lines indicate that, instead of the respective central values, the upper 1 σ bounds of non-perturbative data (MC, open triangles) and perturbative coefficients (PT, filled circles) have been employed. It is remarkable that all lines are approximately constant in the displayed region of β. In addition, we display the empirical minimal term $p_{n_0}\alpha^{n_0+1}$ (dashed green, crosses) in equal units. One sees that the ambiguity in the perturbative series is $\mathcal{O}(20)$ smaller than the condensate. Our explanation why it is not as flat as the other curves is that we only have discrete orders n among which we pick the minimal term.

This is in line with similar observations (for instance, [224]) and important for the QCD sum rules (cf., e.g., [225]). It is reassuring as well: we would not be able to specify a value for the condensate if the difference between non-perturbative plaquette and perturbative resummation were just as big as the ambiguity of the latter.

Finally, we arrive at

$$\langle\frac{\alpha}{\pi}G^2\rangle = 2.68(15)_{n_0}(11)_{\text{PT}}(1)_{\text{MC}}\, r_0^{-4}\,. \tag{5.41}$$

The determination of r_0 entails an error that goes hand in hand with the error of finding the correct energy scale Λ. In our $n_f = 0$ case, we must attribute another error of about 8 percent [226] for scale-setting, which enters with the fourth power and therefore dominates. This leads to

$$\begin{aligned}\langle\frac{\alpha}{\pi}G^2\rangle &= 22.7(74)_{r_0}(12)_{n_0}(10)_{\text{PT}}(1)_{\text{MC}}\left(\Lambda^{\overline{\text{MS}}}_{n_f=0}\right)^4\,,\\ \langle\frac{\alpha}{\pi}G^2\rangle &= 0.0650(210)_{r_0}(36)_{n_0}(28)_{\text{PT}}(3)_{\text{MC}}\,\text{GeV}^4\,.\end{aligned} \tag{5.42}$$

Note that our estimate has the same order of magnitude but otherwise lies significantly above the phenomenological results for the gluon condensate (cf. Sec. 3.10), just as the previous NSPT estimates (cf. Sec. 3.9.2). Regarding the latter, we marvel at their claimed accuracy, given that they neither reach the minimal term nor probe the infinite-volume limit. We add that in [206] a *renormalon-subtracted* scheme is defined whose prescription will lead to a result that is different from Eq. (5.41). This shows once more that the proper treatment of the ambiguity inherent to a factorially growing perturbative series is not without an ambiguity of its own.

6 Summary

From our introduction to the lattice approach (Sec. 1.4), which opens the door to computer simulations in QCD, it became clear that the finite lattice spacing a plays the role of hard regulator. In lattice simulations on a hypercube of size $(La)^4$, the range of accessible momenta k is bounded from above and from below,

$$k \in [k_{\min}, k_{\max}] \sim \left[(La)^{-1}, a^{-1}\right] . \tag{6.1}$$

We learned later on (Sec. 4.2) that a specific choice for the boundary conditions can further reduce $k_{\min}$, but limitations remain.

In this light, the lattice seems to be an unsuited place to search for renormalons (Chap. 3). They designate a particular growth pattern in perturbative expansions at large orders n which causes the series to diverge. Perturbation theory fails to converge because certain momentum regions, for which a perturbative treatment is invalid, intrude into the large-n Feynman diagrams. Particularly severe in asymptotically-free theories such as QCD are the renormalons which arise in the infrared limit $k \to 0$. For momenta as low as $k_{\text{IR}} \sim \Lambda_{\text{QCD}}$, the breakdown of perturbation theory is manifest and cannot be cured.

In hard-regulated theories, renormalons are disguised as divergences in powers of the cutoff (Sec. 3.9.1), which we saw is a^{-1} on the lattice. The power divergences and the renormalon divergences have something in common: both give rise to non-perturbative ambiguities $\sim (\Lambda_{\text{QCD}}/Q)^d$, where Q as a large scale and the dimension d are characteristic for the observable under inspection. The dimension serves to understand and classify renormalon divergence within the OPE framework (Sec. 3.1): the lower d, the stronger the ambiguity due to the renormalon.

A stochastic method called NSPT allows to directly evaluate perturbative coefficients on the lattice (Chap. 2). In the last two decades, different groups have made efforts to reveal the conjectured renormalon in the $d = 4$ plaquette, with inconclusive results so far (Sec. 3.9.2).

In this work, we have made our own attempt to study high-order expansions with NSPT (Chap. 5). In view of the difficulties encountered by the plaquette studies, we first turned to a different observable: the pole mass of a heavy quark or, equivalently, the associated static self-energy (Sec. 3.7). With $d = 1$ it is ideally suited for our purposes because the renormalon ambiguity is maximal.

We calculated the static self-energy with Polyakov loops (Sec. 4.1) up to $\mathcal{O}\left(\alpha^{20}\right)$ on the lattice. Despite conceptual improvements and the massively parallel use of computer

resources (Secs. 2.5 and Appendix A.4), currently feasible simulations are limited to finite lattice volumes with $k_{\min} \gg k_{\mathrm{IR}}$. In other words, the infrared renormalon should be invisible.

We compensated for this deficiency by performing simulations on a wide range of different lattice geometries. To obtain the infinite-volume perturbative coefficients in an extrapolation, one first needs to know the precise form of the FSE. It turns out that there is a one-to-one correspondence between the dimension d of the observable and L^{-d} corrections (Sec. 5.1.3). The latter are accompanied by substantial effects due to a renormalization-group running of the strong coupling α. On accessible lattice volumes, the FSE are of the size of the infinite-volume piece and hence veil the renormalon.

Once FSE and infinite-volume coefficients are disentangled, we find the following: the perturbative series we get for the static self-energy agrees with the leading $d = 1$ renormalon prediction for this observable. With growing order n, the fit is increasingly dominated by the renormalization-group running which we have to insert by hand. A better measure for the significance of our result is therefore given by the normalization of the perturbative series. Due to our very accurate simulation data, we can determine it with an error below 10 percent,

$$N_m^{\mathrm{L}} = 19.1 \pm 1.6\,, \tag{6.2}$$

and it agrees very well with previous (continuum) estimates. The normalization is nothing but the strength of the renormalon pole in the Borel picture (Secs. 3.2 and 5.1.1). Hence, we discover the conjectured renormalon with a certainty of over 10 σ. The pole mass has an intrinsic ambiguity due to the renormalon. It is scheme-independent and reads

$$N_m^{\overline{\mathrm{MS}}} \Lambda_{n_f=0}^{\overline{\mathrm{MS}}} = 153(13)(13)\ \mathrm{MeV}\,, \tag{6.3}$$

which lies within the range of estimates reported by earlier works.

Additional support for our findings not only comes from the agreement of low-order coefficients with literature values, but also from a successful universality check: alternative discretizations of the static action result in differences at low orders (as expected), while the large-order behavior containing the long-distance physics is universal.

With the insights gained in this study, we returned to the plaquette in the second part of this work (Sec. 5.2). The onset of the pole mass renormalon at $n \sim 9$ in the lattice scheme implied the onset of the plaquette renormalon at orders about four times as large and hence a significant increase in scope of the simulations [$\mathcal{O}\left(\alpha^{35}\right)$]. Furthermore, the associated FSE had to be treated correctly. Although fits to our plaquette data as well as to those of another group turned out best with the assumed L^{-4} corrections, we needed an additional L^{-2} term without running coupling to obtain completely satisfactory fit results. Its origin is not entirely understood. A possible explanation are discretization artefacts of the lattice gauge action. Provided that this parametrization is correct, a leading $d = 4$ infrared

renormalon is indeed seen in the extrapolated infinite-volume perturbative series of the plaquette.

Comparing both studies, one realizes that the static self-energy is a much more clean-cut case: firstly, the renormalon ambiguity is so severe that our twenty-loop simulations extend well into the renormalon dominance regime. This allows us to obtain results like the normalization with great confidence. The plaquette renormalon, on the contrary, is so subtle that we can at best see its onset. Secondly, the L^{-1} corrections in the static self-energy clearly dominate, whereas competing effects are conceivable in the $d = 4$ plaquette.

Despite all drawbacks, the perturbative plaquette remains an observable of great interest. The reason is that it allows to extract the gluon condensate from the difference of non-perturbative and perturbative plaquette (Chap. 2). Among all condensates, which in total parametrize QCD vacuum effects, the gluon condensate plays a prominent role (Sec. 3.1). We obtain

$$\begin{aligned}\langle\frac{\alpha}{\pi}G^2\rangle &= 22.7(74)_{r_0}(12)_{n_0}(10)_{\mathrm{PT}}(1)_{\mathrm{MC}}\left(\Lambda^{\overline{\mathrm{MS}}}_{n_f=0}\right)^4\,,\\ \langle\frac{\alpha}{\pi}G^2\rangle &= 0.0650(210)_{r_0}(36)_{n_0}(28)_{\mathrm{PT}}(3)_{\mathrm{MC}}\,\mathrm{GeV}^4\,. \end{aligned} \tag{6.4}$$

Our estimate lies above phenomenological estimates (with problems of their own, see Sec. 3.10), although the discrepancy is less striking in view of our large error (mainly due to scale-setting). Nonetheless, it is encouraging to see that we meet the right order of magnitude with an entirely different approach.

We emphasize that we have consistently worked within pure gauge theory. The qualification we have to make about our results is that they do not stem from full QCD but a world without dynamic quarks. The hope is justified (Secs. 1.1 and 1.3) that what makes QCD special originates in the gluonic sector, and we have picked two observables that are meaningful despite this restriction.

Let us return to where we started the discussion and ask how large a single lattice volume must be to confirm our results without the need of a volume extrapolation. From the study of the static self-energy we learn that we have to consider extents L as large as $\sim e^9 > 8000$ to directly see the onset of renormalon dominance at $n \simeq 9$. It seems that renormalon theory still holds challenges for future computer generations.

Appendix

A.1 Renormalization schemes and the running coupling

Depending on the scheme one chooses for a calculation, one introduces a specific renormalization scale μ. Nonetheless, the result for a physical quantity should be scheme-independent: consider a generic observable $R = R(q, \alpha, m)$, where q represents external momenta, α is the renormalized coupling constant and m denotes the renormalized quark mass (one massive partice considered for simplicity). Scheme-independence leads to the so-called *renormalization-group* (RG) equation

$$\mu \frac{d}{d\mu} R(q, \alpha, m) = 0 = \left(\frac{\partial}{\partial \mu} + \mu \frac{d\alpha}{d\mu} \frac{\partial}{\partial \alpha} + \mu \frac{dm}{d\mu} \frac{\partial}{\partial m} \right) R \,. \tag{A.1}$$

The observable R is then said to be a *RG invariant* and one defines the QCD *β-function* as

$$\beta(\alpha) = \mu \frac{d\alpha}{d\mu} = -2\alpha \left[\beta_0 \frac{\alpha}{4\pi} + \beta_1 \left(\frac{\alpha}{4\pi} \right)^2 + \dots \right], \tag{A.2}$$

and, in analogy, the *mass anomalous dimension*,

$$\gamma(\alpha) = \frac{\mu}{m} \frac{dm}{d\mu} \,, \tag{A.3}$$

as a power series in α with coefficients γ_i. For the β-function, only the first two coefficients are universal for all massless schemes like lattice regularization or the $\overline{\text{MS}}$ scheme [227]:

$$\beta_0 = 11 - \frac{2}{3} n_f, \tag{A.4}$$

$$\beta_1 = 102 - \frac{38}{3} n_f, \tag{A.5}$$

where n_f is the number of quark flavors and we cite the result for the number of colours $N_c = 3$. From β_2 onward the coefficients are scheme-dependent, e.g., in the $\overline{\text{MS}}$ scheme [228]

$$\beta_2^{\overline{\text{MS}}} = \frac{2857}{2} - \frac{5033}{18} n_f + \frac{325}{54} n_f^2. \tag{A.6}$$

Only in this scheme, also β_3 is known [4]. The corresponding value for β_2 in the lattice scheme with Wilson gauge action can be gathered from [229–231]

$$\beta_2^{\text{L}} = \left[-0.00159983232(13) + n_f 0.0000799(4) - n_f^2 0.00000605(2) \right] (4\pi)^3 \,, \tag{A.7}$$

which uses Eq. (A.6) together with the three-loop relation between the two schemes

$$\begin{aligned} \alpha_{\overline{\mathrm{MS}}}(1/a) &= \alpha_{\mathrm{L}} + d_1\alpha_{\mathrm{L}}^2 + d_2\alpha_{\mathrm{L}}^3 + \mathcal{O}\left(\alpha_{\mathrm{L}}^4\right), \\ d_1 &= 5.883591447235371\,, \\ d_2 &= 43.40730288045296\,. \end{aligned} \tag{A.8}$$

The large values for d_1, d_2 hint at a wide separation between the renormalization scales in $\overline{\mathrm{MS}}$ and lattice scheme, respectively. This can be seen explicitly when defining

$$\Lambda_{\mathrm{QCD}} = \lim_{\mu\to\infty} \mu \, \exp\left[-\frac{2\pi}{\beta_0\alpha(\mu)}\right]\left[\frac{\beta_0}{4\pi}\alpha(\mu)\right]^{-\frac{\beta_1}{2\beta_0^2}}, \tag{A.9}$$

as the scale where perturbation theory breaks down. The ratio of Λ_{QCD} evaluated in the two schemes for $n_f = 0$ amounts to

$$\frac{\Lambda_{\mathrm{QCD}}^{\overline{\mathrm{MS}}}}{\Lambda_{\mathrm{QCD}}^{\mathrm{L}}} \approx 28.809\,, \tag{A.10}$$

for the Wilson gauge action on the lattice. In order to relate the coupling strength at two different scales Q^2 and k^2, one can integrate the β-function (A.2). Restricting this to the leading term, it yields

$$\alpha\left(k^2\right) = \frac{\alpha\left(Q^2\right)}{1 - \frac{\beta_0}{4\pi}\alpha\left(Q^2\right)\ln\left(Q^2/k^2\right)}\,. \tag{A.11}$$

We can equally write

$$\alpha\left(k^2\right) = \alpha\left(Q^2\right)\sum_{n=0}^{\infty}\left[\frac{\beta_0}{4\pi}\alpha\left(Q^2\right)\ln\left(\frac{Q^2}{k^2}\right)\right]^n, \tag{A.12}$$

which is known as the *large-β_0* approximation. In this context, the parameter Λ_{QCD} introduced above marks the pole of Eq. (A.11) such that

$$\alpha\left(Q^2\right) \sim \frac{1}{\frac{\beta_0}{4\pi}\ln\left(Q^2/\Lambda_{\mathrm{QCD}}^2\right)}. \tag{A.13}$$

How far one can improve upon the large-β_0 approximation obviously depends upon the knowledge of higher β_i. For instance, this work makes use of the three-loop running

$$\begin{aligned} \alpha\left[1/(La)\right] = \alpha\Bigg\{ & 1 - 2\alpha b_0\lambda \\ & + \alpha\frac{b_1}{b_0}\ln\left[1 - 2\alpha b_0\lambda + \alpha\frac{b_1}{b_0}\ln\left(1 - 2\alpha b_0\lambda\right)\right] \\ & - \left(2\alpha^3 b_0\lambda\right)\left[\frac{b_2}{b_0} - \left(\frac{b_1}{b_0}\right)^2\right]\left(1 - 2\alpha b_0\lambda\right)^{-1}\Bigg\}^{-1}, \end{aligned} \tag{A.14}$$

to relate the couplings $\alpha \equiv \alpha\,(1/a)$ and $\alpha\,[1/(La)]$ in the lattice scheme with each other. For sake of brevity, we have introduced

$$b_i = \frac{\beta_i}{(4\pi)^i}, \tag{A.15}$$

$$\lambda \equiv \ln\,(L)\,, \tag{A.16}$$

such that

$$2\alpha b_0 \lambda = \frac{\beta_0}{4\pi}\alpha \ln\left\{\left[\frac{a^{-1}}{(La)^{-1}}\right]^2\right\}. \tag{A.17}$$

Hence, setting $\beta_1 = \beta_2 = 0$ in Eq. (A.14), we recover the one-loop formula (A.11).

A.2 Stochastic Perturbation Theory

Consider the scalar Euclidian action

$$\begin{aligned} S &= \frac{1}{2}\phi^2 + \frac{1}{4}\lambda\phi^4 \\ &\equiv S_0 + \lambda S_1 \,, \end{aligned} \tag{A.18}$$

where we have split S into a free S_0 and an interaction part λS_1. We not only introduce Langevin time $[\phi \to \phi(t)]$, but also reformulate ϕ as a power series in the coupling λ,

$$\phi(t) = \sum_m \lambda^m \phi^{(m)}(t) \,. \tag{A.19}$$

Eq. (A.19) being a set of linearly independent functions, we plug it into the Langevin equation (2.3) and collect terms of equal power in λ. In doing so, the Langevin equation decomposes into a hierarchical system of equations given by

$$\begin{aligned} \frac{d\phi^{(0)}}{dt} &= -\frac{\partial S_0}{\partial \phi^{(0)}(t)} + \eta(t) = -\phi^{(0)} + \eta(t) \,, \\ \frac{d\phi^{(1)}}{dt} &= -\phi^{(1)}(t) - \left(\phi^{(0)}\right)^3 \,, \\ \frac{d\phi^{(2)}}{dt} &= -\phi^{(2)}(t) - 3\left(\phi^{(0)}\right)^2 \phi^{(1)} \,, \\ \frac{d\phi^{(3)}}{dt} &= \dots \,. \end{aligned} \tag{A.20}$$

Note that the stochastic noise η explicitly apprears only in the equation for $\phi^{(0)}$, whereas all higher $\phi^{(m)}$ are determined by ordinary differential equations.

A.3 Mapping between group and algebra

To shorten the notation, we redefine

$$\begin{aligned} a g_0 A_\mu &\to A_\mu \,, \\ U &= e^{iA} \,. \end{aligned} \tag{A.21}$$

To switch back and forth between algebra and group in terms of a series expansion, we then have the exponential mapping

$$\begin{aligned} U &= \sum_k \frac{i^k \beta^{-\frac{k}{2}}}{k!} A^k = \\ &= \mathbb{1} + i\beta^{-\frac{1}{2}} A^{(1)} + \beta^{-1} \left[iA^{(2)} - \frac{1}{2} \left(A^{(1)} \right)^2 \right] + \\ &\quad + \beta^{-\frac{3}{2}} \left[iA^{(3)} - \frac{1}{2} A^{(1)} A^{(2)} - \frac{1}{2} A^{(2)} A^{(1)} - \frac{i}{3!} \left(A^{(1)} \right)^3 \right] + \dots \,, \end{aligned} \tag{A.22}$$

and the logarithmic mapping

$$\begin{aligned} A &= -i \ln U = -i \ln \left[\mathbb{1} + \sum_{k>0} \beta^{-\frac{k}{2}} U^{(k)} \right] = \\ &= -i\beta^{-\frac{1}{2}} U^{(1)} - i\beta^{-1} \left[U^{(2)} - \frac{1}{2} \left(U^{(1)} \right)^2 \right] + \\ &\quad - i\beta^{-\frac{3}{2}} \left[U^{(3)} - \frac{1}{2} U^{(1)} U^{(2)} - \frac{1}{2} U^{(2)} U^{(1)} + \frac{1}{3} \left(U^{(1)} \right)^3 \right] + \dots \,. \end{aligned} \tag{A.23}$$

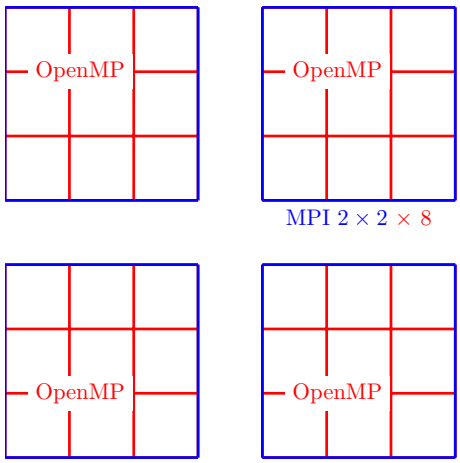

Fig. A.1: Visualization of a hybrid (MPI and OpenMP) simulation using 32 cores. Along two directions, the lattice is split into discrete sublattices which synchronize at given check points in the code by exchanging their common boundaries using MPI, while otherwise working independently. Within one sublattice, the cores attached to one physical memory (8 in the above example) work in parallel, using multi-threading implemented with OpenMP.

A.4 Parallelization of large-scale NSPT simulations

In order to actually implement NSPT on a computer, the most efficient way is to run the code in serial mode, i.e. on a single core. In reality, both the clock rate and the memory available in serial mode can set limits on the scope of the simulation:

- The clock rate restricts the number of floating-point operations that can be carried out per cycle and thus the number of Langevin updates that can be achieved within a given amount of time. Obviously, it is a "soft" limit because in principle one must only wait long enough to complete a Langevin trajectory with sufficient independent measurements to withstand a statistical analysis. In practice, production runs lasting months just to accumulate enough statistics entail heavy drawbacks. Independently of that, Langevin trajectories should be as long as possible to be faithful with the very principle of Stochastic Quantization [Eq. (2.7)].

- Memory restrictions can cause that the intended lattice simulation does not fit on a single machine.

The last point is of practical importance, as a brief calculation illustrates: the memory M_{HDD} needed to store a configuration of SU(N) gauge theory on a symmetric hypercube in D dimensions is determined by the lattice extent L, the maximal loop order M and the size p of a floating point number, which is 8(4) bytes in the case of double (single) machine precision. Compared to M_{HDD}, the memory need during the simulation M_{Run} still increases by a factor I once a higher-order integrator is used ($I = 2$ in the case of our second-order integrator) because one needs to store intermediate results, and also the

matrices containing the gauge rotations increase the memory need. After all, this leads to

$$M_{\mathrm{HDD}} = pDL^D \left[2N^2 (2M) + 2\right] ,$$
$$M_{\mathrm{Run}} = p\,(ID+1)\, L^D \left[2N^2 (2M) + 2\right] . \tag{A.24}$$

For example, we simulated a 32^4 lattice volume with $M = 40$ in double precision requiring $M_{\mathrm{Run}} \sim 101$ Gigabytes. This number clearly exceeds the memory with which serial machines are typically equipped.

If one wants to conduct large-scale simulations, the solution is to *parallelize* the code such that two or more cores team up to complete one simulation. Here, we present two ways to parallelize code that we made use of for this work, either separately or in combination: If the cores all have access to the same physical memory, it is possible to conduct shared-memory multiprocessing using a programming interface such as *OpenMP*, whereas on distributed-memory systems one divides the lattice in several sublattices which are stored (computed) in separate memories (processors). Neighboring sublattices frequently exchange common boundary surfaces among each other via a communication interface such as *MPI*. The downside of both parallelization methods is that they usually make less efficient use of the available cores compared to as many (fictitious) serial runs, which can be quantified in the form of a *load factor* < 100 percent.

In the case of MPI, it is obvious that communication leads to idle time, additionally there may be significant memory overhead in order to deal with the boundary surfaces. OpenMP, by contrast, can be incorporated without major changes. One drawback of OpenMP is that one must protect memory from simultaneous modification by more than one core. Furthermore, only some parts of the code may be suited for parallel execution, and even then idle time can be caused if not all threads terminate in unison. While the performance loss is comparatively small for few cores, OpenMP simulations exhibit poor scaling for a large number of threads.[51] At any rate, the typically very limited number of cores in a shared memory also sets a limit on the scope of pure OpenMP simulations. To get the best of both worlds, one can combine MPI and OpenMP in a *hybrid* simulation with a large number of cores. To get an idea of the overall performance loss due to hybrid parallelization, we analyze a hybrid (MPI and OpenMP) simulation carried out on the SuperMUC cluster at the Leibniz supercomputing centre (LRZ) in Munich. As Fig. A.2 illustrates, the average load factor of the utilized nodes is about 87 percent. From this we conclude that the hybrid code makes quite efficient use of the employed resources.

Let us further examine for the very same example whether the hybrid code maintains a main features of NSPT simulations (cf. Sec. 2.2), namely, that the computational cost scales quadratically with the maximal loop order M. Except for a slight excess, this seems to hold in good approximation, as Fig. A.3 illustrates. Note that the execution time does

[51] We verified this in a $\{28; 35\}$ simulation using 128 threads on a shared-memory partition of the *UV cluster* at the Leibniz supercomputing centre.

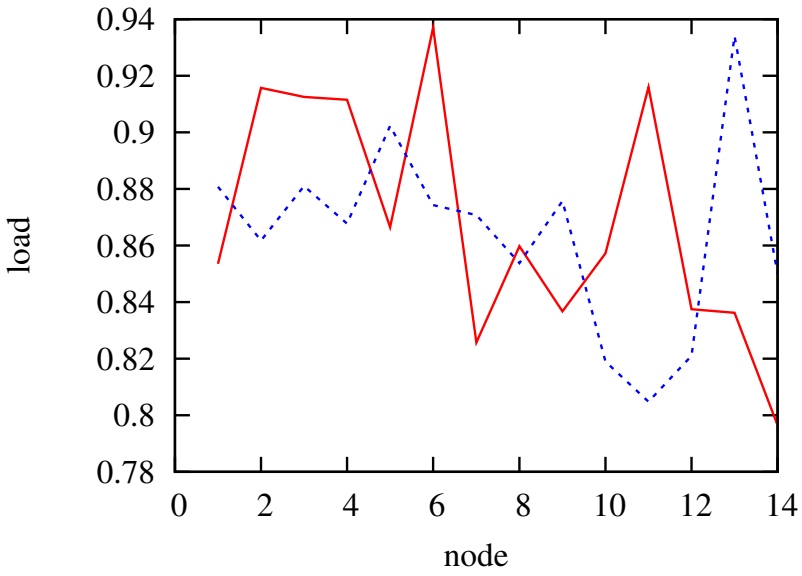

Fig. A.2: The load factor as a fraction of the hypothetical serial performance for each node in two different time intervals. The numbers have been obtained for a hybrid (MPI andOpenMP) $\{28; 35\}$ simulation using a total of 560 cores on the SuperMUC cluster at the LRZ. Averaging over all nodes and both time intervals, the load factor is about 87 percent.

not extrapolate towards zero at low M since M_{Run} was allocated for $M = 35$ already at the beginning of the simulation and hence the MPI exchange between neighboring sublattices always extended to $M = 35$. Assuming that the latter is by far dominant at low orders, one estimates the cost of MPI communication to be ~ 50 seconds per five updates. In view of a total of 450 seconds at $M = 35$, one realizes that MPI communication makes up for just 11 percent of the overall execution time.

In that sense, NSPT code parallelized with MPI becomes more effective (less relative parallelization overhead compared to a fictitious simulation in serial mode) with increasing maximal loop order: while the pure calculation cost exhibits a quadratic scaling behavior, the time needed for MPI exchange scales only linearly with the message size, that is, linearly with the maximal loop order.

We continue analyzing the example run with a crude estimate of the performance loss due to OpenMP. As 11 percent of the overall simulation time is spent on MPI communication and this idle time enters into the calculation of the load factor of 87 percent, we conclude that a mere 2 percent performance loss is due to OpenMP.[52] Its smallness we also attribute to the high maximal loop order of $M = 35$, which causes the pure calculation segment of the code which the various OpenMP threads carry out in parallel with little overhead to be so costly that the simulation spends by far the most time on it.

So while parallel simulations are comparatively efficient at high loop orders M for then they have good performance relative to a serial use of the very same resources, we stress that in general simulations with high M can be even pricier than what the quadratic

[52] In retroperspective, we favor an alternative simulation setup with only 14 MPI subprocesses on a complete node each, using OpenMP with 40 threads.

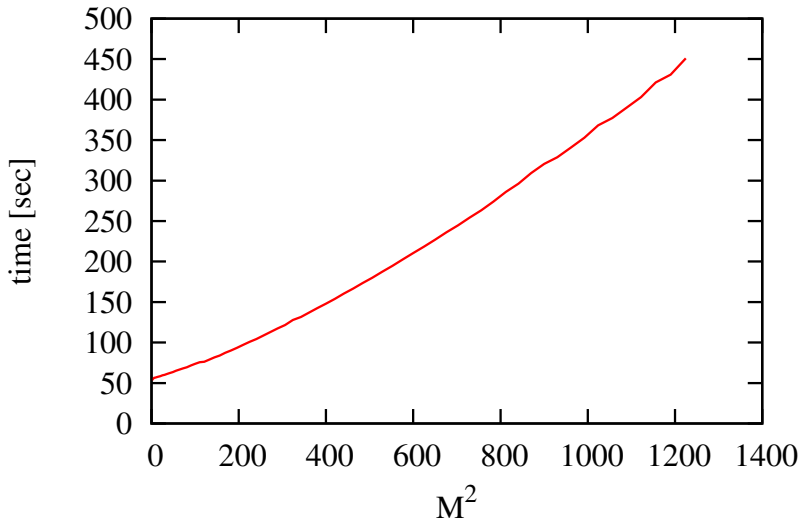

Fig. A.3: Execution time needed to perform 5 Langevin updates as a function of the maximal loop order M^2. In good approximation, the execution time scales quadratically with the maximal loop order. The numbers have been obtained for a hybrid (MPI+OpenMP) $\{28;35\}$ simulation using a total of 560 cores on the SuperMUC cluster. Two lattice directions have been parallelized (MPI 7×4) for a total of 28 MPI subprocesses, two of which each run on one node. In doing so, we wanted to avoid poor scaling behavior when using OpenMP on a complete node (40 cores). Hence, within each MPI subprocess, the work load is distributed among 20 cores in shared-memory fashion using OpenMP.

scaling behavior implies: when monitoring a farm of 200 serial simulations at the LRZ we noted that a small number showed a 50 percent performance increase relative to the bulk part, and this although all cores featured the same clock rate and memory usage was nowhere near threshold. What distinguished the fast cores from the rest was their large cache of 4 Mb as opposed to only 1 Mb. In other words, the simulation was cache-bound. Considering that it requires over 1 kb to store a single 35-loop link expansion and that costly cache reads may be necessary to conduct a link multiplication, we understand why cache size can be the limiting factor for performance.

When fine-tuning the MPI parallelization, the question arises which is the optimal way to divide a lattice into a given number of sublattices. Using MPI,

- our code can be split along either one or two directions,
- the sublattices must be of equal size,
- the minimal edge length is two.

Assume that one wants to distribute a 16^4 lattice on 8 nodes. With the above restrictions, one can choose to either run a sublattice of 2×16^3 (1D MPI) or $4^2 \times 16^2$ (2D MPI) on each node. At first sight, the former has the drawback that, due to the minimal edge length of two, the sublattices are slices with a bad surface-to-volume ratio, whereas the latter has a greater inside unaffected by MPI communication. It turns out, though, that parallelization in two directions requires so *many* MPI exchanges that in almost all cases the performance of parallelizing only one direction was optimal or at least just as good. As

an example, we run a hybrid $\{15; 35\}$ simulation twice: first, we use five nodes to parallelize one direction. Second, we use nine nodes to parallelize two directions on three nodes each. The latter needs 348 seconds on average for five updates, while five nodes require 604 seconds. Rescaling the second number by a factor of 5/9, we see that parallelizing only one direction results in a marginally higher performance. Nonetheless, the MPI functionality in two directions is without alternative if as many nodes as possible have to be employed to be able to fit a large lattice on the machines.

While we saw that the memory needed at run time M_{Run} is a limiting factor in NSPT simulations, the disk space M_{HDD} required to save a configuration does not pose a problem with present-day storages, provided that one discards the configurations on which measurements have been taken and only stores the final state of the Langevin trajectory for continuation. Nonetheless, large parallel simulations turned out to abort frequently when $\mathcal{O}(100)$ sublattices simultaneously wrote their state to disk. Therefore, we were forced to limit the number of parallel input-output operations in these cases.

A.5 Result tables

	$\{\beta_0, \beta_1, \beta_2\}$	$\{\beta_0, \beta_1\}$	$\{\beta_0\}$
χ^2_{red}	1.263	1.290	1.218
$c_0^{(3,0)}$	**2.117274357**	**2.117274357**	**2.117274357**
$c_1^{(3,0)}$	$1.1136(11) \times 10^1$	$1.1136(11) \times 10^1$	$1.1136(11) \times 10^1$
$c_2^{(3,0)}$	$8.610(13) \times 10^1$	$8.610(13) \times 10^1$	$8.597(13) \times 10^1$
$c_3^{(3,0)}$	$7.945(14) \times 10^2$	$7.951(14) \times 10^2$	$7.914(14) \times 10^2$
$c_4^{(3,0)}$	$8.215(26) \times 10^3$	$8.232(26) \times 10^3$	$8.156(26) \times 10^3$
$c_5^{(3,0)}$	$9.322(40) \times 10^4$	$9.361(40) \times 10^4$	$9.203(40) \times 10^4$
$c_6^{(3,0)}$	$1.1533(61) \times 10^6$	$1.1619(61) \times 10^6$	$1.1292(61) \times 10^6$
$c_7^{(3,0)}$	$1.5576(96) \times 10^7$	$1.5760(96) \times 10^7$	$1.5067(94) \times 10^7$
$c_8^{(3,0)}$	$2.304(16) \times 10^8$	$2.345(16) \times 10^8$	$2.194(15) \times 10^8$
$c_9^{(3,0)}$	$3.747(27) \times 10^9$	$3.837(27) \times 10^9$	$3.499(25) \times 10^9$
$c_{10}^{(3,0)}$	$6.702(49) \times 10^{10}$	$6.913(50) \times 10^{10}$	$6.121(46) \times 10^{10}$
$c_{11}^{(3,0)}$	$1.3160(98) \times 10^{12}$	$1.367(10) \times 10^{12}$	$1.1740(89) \times 10^{12}$
$c_{12}^{(3,0)}$	$2.809(24) \times 10^{13}$	$2.939(24) \times 10^{13}$	$2.446(21) \times 10^{13}$
$c_{13}^{(3,0)}$	$6.513(56) \times 10^{14}$	$6.855(58) \times 10^{14}$	$5.537(51) \times 10^{14}$
$c_{14}^{(3,0)}$	$1.628(14) \times 10^{16}$	$1.723(15) \times 10^{16}$	$1.353(13) \times 10^{16}$
$c_{15}^{(3,0)}$	$4.363(38) \times 10^{17}$	$4.641(40) \times 10^{17}$	$3.546(33) \times 10^{17}$
$c_{16}^{(3,0)}$	$1.247(11) \times 10^{19}$	$1.332(11) \times 10^{19}$	$9.925(92) \times 10^{18}$
$c_{17}^{(3,0)}$	$3.785(33) \times 10^{20}$	$4.059(35) \times 10^{20}$	$2.953(28) \times 10^{20}$
$c_{18}^{(3,0)}$	$1.215(11) \times 10^{22}$	$1.308(11) \times 10^{22}$	$9.303(87) \times 10^{21}$
$c_{19}^{(3,0)}$	$4.118(36) \times 10^{23}$	$4.446(38) \times 10^{23}$	$3.094(29) \times 10^{23}$

Table A.1: The goodness of fit and the infinite-volume coefficients of the unsmeared triplet static self-energy for three fits in which different orders of the β-function have been employed. Boldfaced numbers represent DLPT results that have been used in the fit.

	$N_T^{\min} = 8$	$h_{\text{eff}}^{(a)}$	$h_{\text{eff}}^{(b)}$
χ^2_{red}	1.666	0.940	1.033
$c_0^{(3,0)}$	**2.117274357**	**2.117274357**	**2.117274357**
$c_1^{(3,0)}$	$1.1133(10) \times 10^1$	$1.11360(89) \times 10^1$	$1.11442(89) \times 10^1$
$c_2^{(3,0)}$	$8.607(12) \times 10^1$	$8.612(10) \times 10^1$	$8.619(10) \times 10^1$
$c_3^{(3,0)}$	$7.940(12) \times 10^2$	$7.944(10) \times 10^2$	$7.947(10) \times 10^2$
$c_4^{(3,0)}$	$8.201(24) \times 10^3$	$8.233(22) \times 10^3$	$8.231(22) \times 10^3$
$c_5^{(3,0)}$	$9.305(37) \times 10^4$	$9.361(34) \times 10^4$	$9.340(35) \times 10^4$
$c_6^{(3,0)}$	$1.1512(56) \times 10^6$	$1.1606(52) \times 10^6$	$1.1551(53) \times 10^6$
$c_7^{(3,0)}$	$1.5549(88) \times 10^7$	$1.5706(81) \times 10^7$	$1.5589(83) \times 10^7$
$c_8^{(3,0)}$	$2.301(14) \times 10^8$	$2.328(13) \times 10^8$	$2.305(13) \times 10^8$
$c_9^{(3,0)}$	$3.742(24) \times 10^9$	$3.791(23) \times 10^9$	$3.745(23) \times 10^9$
$c_{10}^{(3,0)}$	$6.695(45) \times 10^{10}$	$6.790(41) \times 10^{10}$	$6.695(43) \times 10^{10}$
$c_{11}^{(3,0)}$	$1.3144(89) \times 10^{12}$	$1.3341(82) \times 10^{12}$	$1.3137(85) \times 10^{12}$
$c_{12}^{(3,0)}$	$2.812(20) \times 10^{13}$	$2.850(19) \times 10^{13}$	$2.805(19) \times 10^{13}$
$c_{13}^{(3,0)}$	$6.526(48) \times 10^{14}$	$6.607(44) \times 10^{14}$	$6.490(45) \times 10^{14}$
$c_{14}^{(3,0)}$	$1.632(12) \times 10^{16}$	$1.652(11) \times 10^{16}$	$1.620(11) \times 10^{16}$
$c_{15}^{(3,0)}$	$4.375(33) \times 10^{17}$	$4.426(30) \times 10^{17}$	$4.340(31) \times 10^{17}$
$c_{16}^{(3,0)}$	$1.2506(94) \times 10^{19}$	$1.2650(85) \times 10^{19}$	$1.2401(88) \times 10^{19}$
$c_{17}^{(3,0)}$	$3.796(28) \times 10^{20}$	$3.839(26) \times 10^{20}$	$3.764(27) \times 10^{20}$
$c_{18}^{(3,0)}$	$1.2192(92) \times 10^{22}$	$1.2331(83) \times 10^{22}$	$1.2087(86) \times 10^{22}$
$c_{19}^{(3,0)}$	$4.130(31) \times 10^{23}$	$4.177(28) \times 10^{23}$	$4.094(29) \times 10^{23}$

Table A.2: The goodness of fit and the infinite-volume coefficients for the unsmeared triplet for three fits with different treatments of the N_T dependence. Boldfaced numbers represent DLPT results that have been used in the fit.

$(N_T^{\min}, N_S^{\min})$	$(10, 8)$	$(10, 6)$	$(10, 5)$	$(10, 3)$
χ^2_{red}	1.701	1.431	1.309	1.263
$c_1^{(3,0)}$	11.120(43)	11.124(30)	11.122(19)	11.136(12)
$c_{19}^{(3,0)} \times 10^{-23}$	3.918(96)	3.995(66)	4.108(42)	4.118(41)

Table A.3: χ^2_{red}, $c_1^{(3,0)}$ and $c_{19}^{(3,0)}$ for different $N_S^{\min}$ with the limitation $N_T \geqslant N_S$.

n	$c_n^{(3,0)}$	$c_n^{(3,1/6)}$	$(C_F/C_A)\,c_n^{(8,0)}$	$(C_F/C_A)\,c_n^{(8,1/6)}$
0	**2.117274357**	0.72181(99)	**2.117274357**	0.72181(99)
1	$1.1136(11) \times 10^1$	6.385(10)	$1.1140(12) \times 10^1$	6.387(10)
2	$8.610(13) \times 10^1$	$8.124(12) \times 10^1$	$8.587(14) \times 10^1$	$8.129(12) \times 10^1$
3	$7.945(16) \times 10^2$	$7.670(13) \times 10^2$	$7.917(20) \times 10^2$	$7.682(15) \times 10^2$
4	$8.215(34) \times 10^3$	$8.017(33) \times 10^3$	$8.197(42) \times 10^3$	$8.017(36) \times 10^3$
5	$9.322(59) \times 10^4$	$9.160(59) \times 10^4$	$9.295(76) \times 10^4$	$9.139(64) \times 10^4$
6	$1.153(11) \times 10^6$	$1.138(11) \times 10^6$	$1.144(13) \times 10^6$	$1.134(12) \times 10^6$
7	$1.558(21) \times 10^7$	$1.541(22) \times 10^7$	$1.533(25) \times 10^7$	$1.535(22) \times 10^7$
8	$2.304(43) \times 10^8$	$2.284(45) \times 10^8$	$2.254(51) \times 10^8$	$2.275(45) \times 10^8$
9	$3.747(95) \times 10^9$	$3.717(97) \times 10^9$	$3.64(11) \times 10^9$	$3.703(98) \times 10^9$
10	$6.70(22) \times 10^{10}$	$6.65(22) \times 10^{10}$	$6.49(25) \times 10^{10}$	$6.63(22) \times 10^{10}$
11	$1.316(52) \times 10^{12}$	$1.306(53) \times 10^{12}$	$1.269(59) \times 10^{12}$	$1.303(53) \times 10^{12}$
12	$2.81(13) \times 10^{13}$	$2.79(13) \times 10^{13}$	$2.71(14) \times 10^{13}$	$2.78(13) \times 10^{13}$
13	$6.51(35) \times 10^{14}$	$6.46(35) \times 10^{14}$	$6.29(37) \times 10^{14}$	$6.45(35) \times 10^{14}$
14	$1.628(96) \times 10^{16}$	$1.613(97) \times 10^{16}$	$1.57(10) \times 10^{16}$	$1.614(97) \times 10^{16}$
15	$4.36(28) \times 10^{17}$	$4.32(28) \times 10^{17}$	$4.22(29) \times 10^{17}$	$4.33(28) \times 10^{17}$
16	$1.247(86) \times 10^{19}$	$1.235(86) \times 10^{19}$	$1.206(89) \times 10^{19}$	$1.236(86) \times 10^{19}$
17	$3.78(28) \times 10^{20}$	$3.75(28) \times 10^{20}$	$3.66(28) \times 10^{20}$	$3.75(28) \times 10^{20}$
18	$1.215(93) \times 10^{22}$	$1.204(94) \times 10^{22}$	$1.176(95) \times 10^{22}$	$1.205(94) \times 10^{22}$
19	$4.12(33) \times 10^{23}$	$4.08(33) \times 10^{23}$	$3.99(34) \times 10^{23}$	$4.08(33) \times 10^{23}$

Table A.4: Final result for the perturbative coefficients of the unsmeared and smeared static self-energies in fundamental and adjoint representations. Boldfaced numbers represent DLPT results that have been used in the fit.

n	p_n	f_n	h_n
0	$\mathbf{4\pi/3}$	$--$	$--$
1	$\mathbf{5.355009}$	0.22(57)	$--$
2	$\mathbf{2.719832 \times 10^1}$	$-1.7(59)$	0.19(17)
3	$1.80962(11) \times 10^2$	$4.8(66) \times 10^1$	$-2.6(21)$
4	$1.38315(14) \times 10^3$	$-1.0(90) \times 10^2$	$-3.9(33) \times 10^1$
5	$1.15150(16) \times 10^4$	$8(10) \times 10^3$	$-7.0(38) \times 10^2$
6	$1.01625(19) \times 10^5$	$7(120) \times 10^3$	$-1.07(45) \times 10^4$
7	$9.3582(21) \times 10^5$	$5(15) \times 10^5$	$-1.54(51) \times 10^5$
8	$8.9015(21) \times 10^6$	$2(18) \times 10^6$	$-2.15(51) \times 10^6$
9	$8.6854(20) \times 10^7$	$5(21) \times 10^7$	$-3.13(47) \times 10^7$
10	$8.6521(33) \times 10^8$	$1.4(27) \times 10^9$	$-5.38(81) \times 10^8$
11	$8.7702(57) \times 10^9$	$-7(43) \times 10^9$	$-8.2(14) \times 10^9$
12	$9.004(13) \times 10^{10}$	$2.9(77) \times 10^{11}$	$-1.25(22) \times 10^{11}$
13	$9.368(18) \times 10^{11}$	$1(13) \times 10^{12}$	$-1.95(34) \times 10^{12}$
14	$9.850(25) \times 10^{12}$	$3(22) \times 10^{13}$	$-3.09(53) \times 10^{13}$
15	$1.0455(33) \times 10^{14}$	$2(38) \times 10^{14}$	$-4.76(84) \times 10^{14}$
16	$1.1186(46) \times 10^{15}$	$9(64) \times 10^{15}$	$-7.5(14) \times 10^{15}$
17	$1.2052(63) \times 10^{16}$	$-5(1100) \times 10^{15}$	$-1.18(23) \times 10^{17}$
18	$1.3076(85) \times 10^{17}$	$1(19) \times 10^{18}$	$-1.86(39) \times 10^{18}$
19	$1.428(12) \times 10^{18}$	$1(34) \times 10^{19}$	$-2.93(69) \times 10^{19}$
20	$1.567(16) \times 10^{19}$	$3(60) \times 10^{20}$	$-4.8(13) \times 10^{20}$
21	$1.732(21) \times 10^{20}$	$3(110) \times 10^{21}$	$-8.0(24) \times 10^{21}$
22	$1.925(28) \times 10^{21}$	$1(20) \times 10^{23}$	$-1.39(49) \times 10^{23}$
23	$2.150(38) \times 10^{22}$	$2(36) \times 10^{24}$	$-2.6(10) \times 10^{24}$
24	$2.414(54) \times 10^{23}$	$4(65) \times 10^{25}$	$-5.0(22) \times 10^{25}$
25	$2.724(77) \times 10^{24}$	$1(12) \times 10^{27}$	$-1.03(50) \times 10^{27}$
26	$3.09(11) \times 10^{25}$	$2(23) \times 10^{28}$	$-2.3(12) \times 10^{28}$
27	$3.53(17) \times 10^{26}$	$6(43) \times 10^{29}$	$-5.3(29) \times 10^{29}$
28	$4.08(25) \times 10^{27}$	$1.0(83) \times 10^{31}$	$-1.30(72) \times 10^{31}$
29	$4.84(37) \times 10^{28}$	$4(16) \times 10^{32}$	$-3.3(19) \times 10^{32}$
30	$5.95(60) \times 10^{29}$	$8(32) \times 10^{33}$	$-8.8(51) \times 10^{33}$
31	$7.66(93) \times 10^{30}$	$2.6(63) \times 10^{35}$	$-2.4(14) \times 10^{35}$
32	$1.03(14) \times 10^{32}$	$7(13) \times 10^{36}$	$-7.0(41) \times 10^{36}$
33	$1.48(27) \times 10^{33}$	$2.0(27) \times 10^{38}$	$-2.0(12) \times 10^{38}$
34	$2.56(45) \times 10^{34}$	$6.3(59) \times 10^{39}$	$-6.2(37) \times 10^{39}$

Table A.5: Final result for the perturbative plaquette, FSE and artefact coefficients. Boldfaced DLPT numbers, e.g., from [131].

n	$p_n/(np_{n-1})$	$f_n/(nf_{n-1})$	$h_n/(nh_{n-1})$
1	**1.2784**	$--$	$--$
2	**2.5395**	$-4(10)$	$--$
3	*2.21781(13)*	$-9.5(29)$	$-4.7(61)$
4	1.91083(22)	$-0.5(44)$	3.8(39)
5	1.66504(27)	$-2(13) \times 10^1$	3.6(29)
6	1.47090(31)	0.1(26)	2.5(14)
7	1.31551(34)	$1(20) \times 10^1$	2.06(86)
8	1.18900(34)	0.6(63)	1.75(52)
9	1.08414(32)	2(23)	1.62(30)
10	0.99616(42)	3(19)	1.72(27)
11	0.92151(65)	$-0.4(23)$	1.38(25)
12	0.8555(13)	$-3.7(17)$	1.27(22)
13	0.8004(18)	0.3(41)	1.20(19)
14	0.7510(23)	2(36)	1.13(17)
15	0.7076(27)	0.4(12)	1.03(14)
16	0.6687(32)	3(100)	0.98(12)
17	0.6338(39)	$-3(690) \times 10^{-2}$	0.93(11)
18	0.6027(46)	$-1(290) \times 10^1$	0.873(99)
19	0.5749(54)	0.4(20)	0.830(90)
20	0.5487(62)	2(91)	0.816(86)
21	0.5262(70)	0.4(22)	0.795(79)
22	0.5051(78)	2(99)	0.792(72)
23	0.4857(87)	0.7(25)	0.799(64)
24	0.4677(98)	1(35)	0.812(55)
25	0.451(11)	1(29)	0.830(45)
26	0.436(13)	0.7(17)	0.848(34)
27	0.423(15)	1(22)	0.864(24)
28	0.413(18)	0.7(10)	0.874(17)
29	0.409(21)	1(15)	0.882(11)
30	0.410(26)	0.7(53)	0.8867(79)
31	0.415(27)	1.1(71)	0.8899(58)
32	0.422(25)	0.8(33)	0.8915(46)
33	0.432(44)	0.9(27)	0.8923(39)
34	0.510(53)	0.9(17)	0.8922(36)

Table A.6: Ratios of infinite-volume plaquette coefficients, FSE and artefact contributions. Boldfaced (italicized) ratios rely (in part) on DLPT results from [131].

Bibliography

[1] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. **B716**, 1 (2012), arXiv:1207.7214 [hep-ex].

[2] T. Aoyama, M. Hayakawa, T. Kinoshita and M. Nio, Phys. Rev. Lett. **109**, 111807 (2012), arXiv:1205.5368 [hep-ph].

[3] F. J. Dyson, Phys. Rev. **85**, 631 (1952).

[4] T. van Ritbergen, J. A. M. Vermaseren and S. A. Larin, Phys. Lett. **B400**, 379 (1997), arXiv:hep-ph/9701390.

[5] P. A. Baikov, K. G. Chetyrkin and J. H. Kühn, Phys. Rev. Lett. **104**, 132004 (2010), arXiv:1001.3606 [hep-ph].

[6] K. Melnikov and T. v. Ritbergen, Phys. Lett. **B482**, 99 (2000), arXiv:hep-ph/9912391.

[7] K. G. Chetyrkin and M. Steinhauser, Nucl. Phys. **B573**, 617 (2000), arXiv:hep-ph/9911434.

[8] A. A. Penin and M. Steinhauser, Phys. Lett. **B538**, 335 (2002), arXiv:hep-ph/0204290 [hep-ph].

[9] F. David, Nucl. Phys. **B209**, 433 (1982).

[10] C. M. Bender and T. T. Wu, Phys. Rev. **D7**, 1620 (1973).

[11] F. Di Renzo, E. Onofri, G. Marchesini and P. Marenzoni, Nucl. Phys. **B426**, 675 (1994), arXiv:hep-lat/9405019.

[12] K. G. Wilson, Phys. Rev. **D10**, 2445 (1974).

[13] K. Nakamura *et al.* (Particle Data Group), J. Phys. **G37**, 075021 (2010).

[14] S. Bethke, Eur. Phys. J. **C64**, 689 (2009), arXiv:0908.1135 [hep-ph].

[15] D. J. Gross and F. Wilczek, Phys. Rev. Lett. **30**, 1343 (1973).

[16] H. D. Politzer, Phys. Rev. Lett. **30**, 1346 (1973).

[17] K. Symanzik, Nucl. Phys. **B226**, 187 (1983).

[18] P. Weisz, Nucl. Phys. **B212**, 1 (1983).

[19] P. Weisz and R. Wohlert, Nucl. Phys. **B236**, 397 (1984).

[20] M. Lüscher and P. Weisz, Commun. Math. Phys. **97**, 59 (1985).

[21] H. B. Nielsen and M. Ninomiya, Nucl. Phys. **B185**, 20 (1981).

[22] C. Gattringer and C. B. Lang, Lect. Notes Phys. **788**, 1 (2010).

[23] Z. Fodor and C. Hoelbling, Rev. Mod. Phys. **84**, 449 (2012), arXiv:1203.4789 [hep-lat].

[24] P. Langevin, Comptes Rendues **146**, 530 (1908).

[25] G. Parisi and Y. S. Wu, Sci. Sin. **24**, 483 (1981).

[26] B. Svetitsky, Phys. Rev. **D37**, 2484 (1988).

[27] J. Ambjørn, R. Loll, W. Westra and S. Zohren, Phys. Lett. **B680**, 359 (2009), arXiv:0908.4224 [hep-th].

[28] G. Aarts, Phys. Rev. Lett. **102**, 131601 (2009), arXiv:0810.2089 [hep-lat].

[29] A. A. Migdal, Sov. Phys. Usp. **29**, 389 (1986).

[30] P. H. Damgaard and H. Hüffel, Phys. Rept. **152**, 227 (1987).

[31] M. Namiki, I. Ohba, K. Okano, Y. Yamanaka, A. K. Kapoor, H. Nakazato and S. Tanaka, Lect. Notes Phys. **M9**, 1 (1992).

[32] E. Floratos and J. Iliopoulos, Nucl. Phys. **B214**, 392 (1983).

[33] W. Grimus and H. Hüffel, Z. Phys. **C18**, 129 (1983).

[34] I. T. Drummond, S. Duane and R. R. Horgan, Nucl. Phys. **B220**, 119 (1983).

[35] G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky and K. G. Wilson, Phys. Rev. **D32**, 2736 (1985).

[36] G. S. Bali, C. Bauer, A. Pineda and C. Torrero, in preparation (2012).

[37] M. Beneke, Phys. Rept. **317**, 1 (1999), arXiv:hep-ph/9807443.

[38] S. Capitani, Phys. Rept. **382**, 113 (2003), arXiv:hep-lat/0211036.

[39] D. Zwanziger, Nucl. Phys. **B192**, 259 (1981).

[40] L. Baulieu and D. Zwanziger, Nucl. Phys. **B193**, 163 (1981).

[41] F. Di Renzo, E.-M. Ilgenfritz, H. Perlt, A. Schiller and C. Torrero, Nucl. Phys. **B831**, 262 (2010), arXiv:0912.4152 [hep-lat].

[42] F. Di Renzo, E.-M. Ilgenfritz, H. Perlt, A. Schiller and C. Torrero, Nucl. Phys. **B842**, 122 (2011), arXiv:1008.2617 [hep-lat].

[43] F. Di Renzo, E. Onofri and G. Marchesini, Nucl. Phys. **B457**, 202 (1995), arXiv:hep-th/9502095.

[44] U. M. Heller and F. Karsch, Nucl. Phys. **B251**, 254 (1985).

[45] E. V. Shuryak, Phys. Rept. **115**, 151 (1984).

[46] P. Colangelo and A. Khodjamirian, in *At the Frontier of Particle Physics. Handbook of QCD*, Vol. 3, edited by M. A. Shifman (World Scientific, Singapore, New Jersey, 2002), arXiv:hep-ph/0010175.

[47] K. G. Wilson, Phys. Rev. **179**, 1499 (1969).

[48] R. A. Brandt, Fortsch. Phys. **18**, 249 (1970).

[49] K. Symanzik, Commun. Math. Phys. **23**, 49 (1971).

[50] J. Callan, Curtis G., Phys. Rev. **D5**, 3202 (1972).

[51] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. **B147**, 385 (1979).

[52] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. **B147**, 448 (1979).

[53] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. **B147**, 519 (1979).

[54] G. 't Hooft, Phys. Rev. **D14**, 3432 (1976).

[55] J. Callan, Curtis G., R. F. Dashen and D. J. Gross, Phys. Rev. **D17**, 2717 (1978).

[56] D. B. Leinweber, Annals Phys. **254**, 328 (1997), arXiv:nucl-th/9510051.

[57] S. Narison, Nucl. Phys. Proc. Suppl. **207-208**, 315 (2010), arXiv:1010.1959 [hep-ph].

[58] S. J. Brodsky, C. D. Roberts, R. Shrock and P. C. Tandy, Phys. Rev. **C82**, 022201 (2010), arXiv:1005.4610 [nucl-th].

[59] S. J. Brodsky, C. D. Roberts, R. Shrock and P. C. Tandy, Phys. Rev. **C85**, 065202 (2012), arXiv:1202.2376 [nucl-th].

[60] S. Narison, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **17**, 1 (2002), arXiv:hep-ph/0205006.

[61] J. C. Collins, A. Duncan and S. D. Joglekar, Phys. Rev. **D16**, 438 (1977).

[62] R. Tarrach, Nucl. Phys. **B196**, 45 (1982).

[63] M. Gell-Mann, R. J. Oakes and B. Renner, Phys. Rev. **175**, 2195 (1968).

[64] H. Fukaya *et al.* (JLQCD collaboration), Phys. Rev. Lett. **104**, 122002 (2010), arXiv:0911.5555 [hep-lat].

[65] A. H. Mueller, Nucl. Phys. **B250**, 327 (1985).

[66] F. David, Nucl. Phys. **B234**, 237 (1984).

[67] F. David, Nucl. Phys. **B263**, 637 (1986).

[68] V. I. Zakharov, Nucl. Phys. **B385**, 452 (1992).

[69] M. Beneke and V. I. Zakharov, Phys. Rev. Lett. **69**, 2472 (1992).

[70] M. Beneke, Nucl. Phys. **B405**, 407 (1993).

[71] S. L. Adler, Phys. Rev. **D10**, 3714 (1974).

[72] P. A. Baikov, K. G. Chetyrkin and J. H. Kühn, Phys. Rev. Lett. **101**, 012002 (2008), arXiv:0801.1821 [hep-ph].

[73] D. J. Gross and A. Neveu, Phys. Rev. **D10**, 3235 (1974).

[74] B. E. Lautrup, Phys. Lett. **B69**, 109 (1977).

[75] G. 't Hooft, in *Proc. Int. School: The whys of subnuclear physics*, edited by A. Zichichi, Erice 1977 (Plenum, New York, 1978).

[76] I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, Phys. Rev. **D50**, 2234 (1994), arXiv:hep-ph/9402360.

[77] V. I. Zakharov, Nucl. Phys. **B377**, 501 (1992).

[78] M. Neubert, Phys. Rev. **D51**, 5924 (1995), arXiv:hep-ph/9412265.

[79] K. G. Chetyrkin, V. Spiridonov and S. Gorishnii, Phys. Lett. **B160**, 149 (1985).

[80] L. R. Surguladze and F. V. Tkachov, Nucl. Phys. **B331**, 35 (1990).

[81] M. Beneke, Nucl. Phys. **B405**, 424 (1993).

[82] D. J. Broadhurst, Z. Phys. **C58**, 339 (1993).

[83] M. Beneke and V. A. Smirnov, Nucl. Phys. **B472**, 529 (1996), arXiv:hep-ph/9510437.

[84] G. Parisi, Phys. Lett. **B76**, 65 (1978).

[85] A. I. Vainshtein and V. I. Zakharov, Phys. Rev. Lett. **73**, 1207 (1994), arXiv:hep-ph/9404248.

[86] G. Di Cecio and G. Paffuti, Int. J. Mod. Phys. **A10**, 1449 (1995), arXiv:hep-th/9406006.

[87] D. Diakonov, Nucl. Phys. Proc. Suppl. **195**, 5 (2009), arXiv:0906.2456 [hep-ph].

[88] L. N. Lipatov, JETP Lett. **25**, 104 (1977).

[89] A. A. Belavin, A. M. Polyakov, A. S. Schwartz and Y. Tyupkin, Phys. Lett. **B59**, 85 (1975).

[90] G. 't Hooft, Phys. Rev. Lett. **37**, 8 (1976).

[91] I. I. Balitsky, Phys. Lett. **B273**, 282 (1991).

[92] E. B. Bogomolny and V. A. Fateyev, Phys. Lett. **B71**, 93 (1977).

[93] J. C. Le Guillou and J. Zinn-Justin, *Large order behavior of perturbation theory*, edited by J. C. Le Guillou and J. Zinn-Justin (North-Holland, Amsterdam, Netherlands, 1990).

[94] A. X. El-Khadra and M. Luke, Ann. Rev. Nucl. Part. Sci. **52**, 201 (2002), arXiv:hep-ph/0208114.

[95] R. Tarrach, Nucl. Phys. **B183**, 384 (1981).

[96] D. Johnston, preprint LPTHE Orsay 86/49 (1986).

[97] J. C. Breckenridge, M. J. Lavelle and T. G. Steele, Z. Phys. **C65**, 155 (1995), arXiv:hep-th/9407028.

[98] A. S. Kronfeld, Phys. Rev. **D58**, 051501 (1998), arXiv:hep-ph/9805215.

[99] N. Gray, D. J. Broadhurst, W. Grafe and K. Schilcher, Z. Phys. **C48**, 673 (1990).

[100] A. F. Falk, preprint JHU-TIPAC-20005 (2000), arXiv:hep-ph/0007339.

[101] M. Beneke and V. M. Braun, Nucl. Phys. **B426**, 301 (1994), arXiv:hep-ph/9402364.

[102] M. E. Luke, Phys. Lett. **B252**, 447 (1990).

[103] E. Eichten and B. R. Hill, Phys. Lett. **B234**, 511 (1990).

[104] H. Georgi, Phys. Lett. **B240**, 447 (1990).

[105] A. Pineda, *Heavy quarkonium and nonrelativistic effective field theories*, Ph.D. thesis, Universitat de Barcelona (1998).

[106] A. F. Falk, M. Neubert and M. E. Luke, Nucl. Phys. **B388**, 363 (1992), arXiv:hep-ph/9204229.

[107] I. I. Y. Bigi, M. A. Shifman and N. Uraltsev, Ann. Rev. Nucl. Part. Sci. **47**, 591 (1997), arXiv:hep-ph/9703290.

[108] I. I. Y. Bigi and N. G. Uraltsev, Phys. Lett. **B321**, 412 (1994), arXiv:hep-ph/9311337.

[109] I. I. Y. Bigi, N. G. Uraltsev and A. I. Vainshtein, Phys. Lett. **B293**, 430 (1992), arXiv:hep-ph/9207214.

[110] P. Ball, M. Beneke and V. M. Braun, Nucl. Phys. **B452**, 563 (1995), arXiv:hep-ph/9502300.

[111] M. B. Voloshin, Int. J. Mod. Phys. **A10**, 2865 (1995), arXiv:hep-ph/9502224.

[112] M. Beneke, V. M. Braun and V. I. Zakharov, Phys. Rev. Lett. **73**, 3058 (1994), arXiv:hep-ph/9405304.

[113] M. E. Luke, A. V. Manohar and M. J. Savage, Phys. Rev. **D51**, 4924 (1995), arXiv:hep-ph/9407407.

[114] M. Neubert and C. T. Sachrajda, Nucl. Phys. **B438**, 235 (1995), arXiv:hep-ph/9407394.

[115] E. Eichten, Nucl. Phys. Proc. Suppl. **4**, 170 (1988).

[116] L. Maiani, G. Martinelli and C. T. Sachrajda, Nucl. Phys. **B368**, 281 (1992).

[117] G. Martinelli, M. Neubert and C. T. Sachrajda, Nucl. Phys. **B461**, 238 (1996), arXiv:hep-ph/9504217.

[118] E. Eichten and B. R. Hill, Phys. Lett. **B240**, 193 (1990).

[119] P. Boucaud, L. C. Lung and O. Pène, Phys. Rev. **D40**, 1529 (1989).

[120] G. Martinelli and C. T. Sachrajda, Phys. Lett. **B354**, 423 (1995), arXiv:hep-ph/9502352.

[121] M. Crisafulli, V. Giménez, G. Martinelli and C. T. Sachrajda, Nucl. Phys. **B457**, 594 (1995), arXiv:hep-ph/9506210.

[122] V. Giménez, G. Martinelli and C. T. Sachrajda, Nucl. Phys. **B486**, 227 (1997), arXiv:hep-lat/9607055.

[123] V. Giménez, G. Martinelli and C. T. Sachrajda, Nucl. Phys. Proc. Suppl. **53**, 365 (1997), arXiv:hep-lat/9608039.

[124] V. Giménez, G. Martinelli and C. T. Sachrajda, Phys. Lett. **B393**, 124 (1997), arXiv:hep-lat/9607018.

[125] V. Giménez, L. Giusti, G. Martinelli and F. Rapuano, JHEP **0003**, 018 (2000), arXiv:hep-lat/0002007.

[126] J. E. Mandula, G. Zweig, and J. Govaerts, Nucl. Phys. **B228**, 109 (1983).

[127] M. A. Nobes, H. D. Trottier, G. P. Lepage and Q. Mason, Nucl. Phys. Proc. Suppl. **106**, 838 (2002), arXiv:hep-lat/0110051.

[128] M. Creutz, Phys. Rev. Lett. **43**, 553 (1979).

[129] A. Di Giacomo and G. Rossi, Phys. Lett. **B100**, 481 (1981).

[130] B. Allés, M. Campostrini, A. Feo and H. Panagopoulos, Phys. Lett. **B324**, 433 (1994), arXiv:hep-lat/9306001.

[131] B. Allés, A. Feo and H. Panagopoulos, Phys. Lett. **B426**, 361 (1998), arXiv:hep-lat/9801003.

[132] P. E. L. Rakow, PoS **LAT2005**, 284 (2006), arXiv:hep-lat/0510046.

[133] F. Di Renzo and L. Scorzato, JHEP **0110**, 038 (2001), arXiv:hep-lat/0011067.

[134] E.-M. Ilgenfritz, Y. Nakamura, H. Perlt, P. E. L. Rakow, G. Schierholz and A. Schiller, PoS **LAT2009**, 236 (2009), arXiv:0910.2795 [hep-lat].

[135] R. Horsley, G. Hotzel, E.-M. Ilgenfritz, R. Millo, H. Perlt, P. E. L. Rakow, Y. Nakamura, G. Schierholz and A. Schiller, Phys. Rev. **D86**, 054502 (2012), arXiv:1205.1659 [hep-lat].

[136] G. Burgio, F. Di Renzo, G. Marchesini and E. Onofri, Phys. Lett. **B422**, 219 (1998), arXiv:hep-ph/9706209.

[137] G. Burgio, F. Di Renzo, G. Marchesini, E. Onofri, M. Pepe and L. Scorzato, Nucl. Phys. Proc. Suppl. **63**, 808 (1998), arXiv:hep-lat/9709106.

[138] R. Horsley, P. E. L. Rakow and G. Schierholz, Nucl. Phys. Proc. Suppl. **106**, 870 (2002), arXiv:hep-lat/0110210.

[139] S. Narison and V. I. Zakharov, Phys. Lett. **B679**, 355 (2009), arXiv:0906.4312 [hep-ph].

[140] Y. Meurice, Phys. Rev. **D74**, 096005 (2006), arXiv:hep-lat/0609005.

[141] S. Narison, Phys. Lett. **B361**, 121 (1995), arXiv:hep-ph/9504334.

[142] J. Bordes, V. Giménez and J. A. Penarrocha, Phys. Lett. **B201**, 365 (1988).

[143] M. B. Causse and G. Mennessier, Z. Phys. **C47**, 611 (1990).

[144] S. Narison, Phys. Lett. **B387**, 162 (1996), arXiv:hep-ph/9512348.

[145] B. L. Ioffe, Prog. Part. Nucl. Phys. **56**, 232 (2006), arXiv:hep-ph/0502148.

[146] S. Narison, Phys. Lett. **B673**, 30 (2009), arXiv:0901.3823 [hep-ph].

[147] K. Zyablyuk, JHEP **0301**, 081 (2003), arXiv:hep-ph/0210103.

[148] A. Samsonov, preprint (2004), arXiv:hep-ph/0407199.

[149] B. L. Ioffe and K. N. Zyablyuk, Eur. Phys. J. **C27**, 229 (2003), arXiv:hep-ph/0207183.

[150] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).

[151] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[152] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).

[153] M. Glück, E. Reya and A. Vogt, Z. Phys. **C53**, 651 (1992).

[154] B. L. Ioffe and A. G. Oganesian, Nucl. Phys. **A714**, 145 (2003), arXiv:hep-ph/0208192.

[155] C. M. Bender and T. T. Wu, Phys. Rev. Lett. **27**, 461 (1971).

[156] E. Brézin, J. C. Le Guillou and J. Zinn-Justin, Phys. Rev. **D15**, 1544 (1977).

[157] E. Brézin, J. Le Guillou and J. Zinn-Justin, Phys. Rev. **D15**, 1558 (1977).

[158] Y. Frishman and S. Yankielowicz, Phys. Rev. **D19**, 540 (1979).

[159] I. Affleck, Nucl. Phys. **B191**, 429 (1981).

[160] M. C. Bergère and F. David, Phys. Lett. **B135**, 412 (1984).

[161] G. Parisi, Phys. Rept. **49**, 215 (1979).

[162] G. Parisi, Phys. Lett. **B66**, 167 (1977).

[163] A. M. Polyakov, Phys. Lett. **B59**, 79 (1975).

[164] A. R. Zhitnitsky, Phys. Rev. **D53**, 5821 (1996), arXiv:hep-ph/9510366.

[165] D. J. Gross and V. Periwal, Phys. Rev. Lett. **60**, 2105 (1988).

[166] J.-P. Eckmann, J. Magnen and R. Sénéor, Commun. Math. Phys. **39**, 251 (1975).

[167] J. Magnen and R. Sénéor, Commun. Math. Phys. **56**, 237 (1977).

[168] C. Itzykson, G. Parisi and J. B. Zuber, Phys. Rev. Lett. **38**, 306 (1977).

[169] C. A. Hurst, Proc. Cambridge Phil. Soc. **48**, 625 (1952).

[170] W. E. Thirring, Helv. Phys. Acta **26**, 33 (1953).

[171] A. Peterman, Helv. Phys. Acta **26**, 291 (1953).

[172] M. Fry, Phys. Lett. **B80**, 65 (1978).

[173] P. Renuard, preprint École Polytechnique (Palaiseau) A 247.1076 (1976).

[174] C. Itzykson, G. Parisi and J. B. Zuber, Phys. Rev. **D16**, 996 (1977).

[175] R. Balian, C. Itzykson, G. Parisi and J. B. Zuber, Phys. Rev. **D17**, 1041 (1978).

[176] J. Fischer, Int. J. Mod. Phys. **A12**, 3625 (1997), arXiv:hep-ph/9704351.

[177] L. D. McLerran and B. Svetitsky, Phys. Rev. **D24**, 450 (1981).

[178] G. 't Hooft, Nucl. Phys. **B153**, 141 (1979).

[179] G. 't Hooft, Commun. Math. Phys. **81**, 267 (1981).

[180] G. Parisi, in *Progress in Gauge Field Theory*, edited by G. 't Hooft *et al.*, Cargèse Lectures 1983 (Plenum, New York, 1984).

[181] M. Lüscher and P. Weisz, Nucl. Phys. **B266**, 309 (1986).

[182] A. González-Arroyo and C. P. Korthals Altes, Nucl. Phys. **B311**, 433 (1988).

[183] P. van Baal, Commun. Math. Phys. **92**, 1 (1983).

[184] R. Wohlert, preprint DESY 87/069 (1987).

[185] G. 't Hooft, Nucl. Phys. **B79**, 276 (1974).

[186] M. Srednicki and L. Susskind, Nucl. Phys. **B179**, 239 (1981).

[187] G. 't Hooft, Nucl. Phys. **B138**, 1 (1978).

[188] J. Groeneveld, J. Jurkiewicz and C. P. Korthals Altes, Phys. Scripta **23**, 1022 (1981).

[189] T. A. DeGrand and D. Toussaint, Phys. Rev. **D25**, 526 (1982).

[190] T. G. Kovacs and E. T. Tomboulis, Phys. Rev. Lett. **85**, 704 (2000), arXiv:hep-lat/0002004.

[191] A. González-Arroyo and M. Okawa, Phys. Lett. **B120**, 174 (1983).

[192] T. Eguchi and H. Kawai, Phys. Rev. Lett. **48**, 1063 (1982).

[193] G. de Divitiis, R. Frezzotti, M. Guagnelli and R. Petronzio, Nucl. Phys. **B422**, 382 (1994), arXiv:hep-lat/9312085.

[194] G. de Divitiis, R. Frezzotti, M. Guagnelli and R. Petronzio, Nucl. Phys. **B433**, 390 (1995), arXiv:hep-lat/9407028.

[195] H. Perlt and A. Schiller, private communication (2011).

[196] F. Di Renzo and L. Scorzato, JHEP **0410**, 073 (2004), arXiv:hep-lat/0410010.

[197] H. D. Trottier, N. H. Shakespeare, G. P. Lepage and P. B. Mackenzie, Phys. Rev. **D65**, 094502 (2002), arXiv:hep-lat/0111028.

[198] T. Blum, C. E. DeTar, S. A. Gottlieb, K. Rummukainen, U. M. Heller, J. E. Hetrick, D. Toussaint, R. L. Sugar and M. Wingate, Phys. Rev. **D55**, 1133 (1997), arXiv:hep-lat/9609036.

[199] G. P. Lepage, Phys. Rev. **D59**, 074502 (1999), arXiv:hep-lat/9809157.

[200] A. Hasenfratz and F. Knechtli, Phys. Rev. **D64**, 034504 (2001), arXiv:hep-lat/0103029.

[201] M. Falcioni, M. L. Paciello, G. Parisi and B. Taglienti, Nucl. Phys. **B251**, 624 (1985).

[202] M. Albanese *et al.* (APE), Phys. Lett. **B192**, 163 (1987).

[203] C. Morningstar and M. J. Peardon, Phys. Rev. **D69**, 054501 (2004), arXiv:hep-lat/0311018.

[204] C. Bauer, G. S. Bali and A. Pineda, Phys. Rev. Lett. **108**, 242002 (2012), arXiv:1111.3946 [hep-ph].

[205] M. Beneke, Phys. Lett. **B344**, 341 (1995), arXiv:hep-ph/9408380.

[206] A. Pineda, JHEP **0106**, 022 (2001), arXiv:hep-ph/0105008.

[207] A. H. Hoang, M. C. Smith, T. Stelzer and S. Willenbrock, Phys. Rev. **D59**, 114014 (1999), arXiv:hep-ph/9804227.

[208] M. Beneke, Phys. Lett. **B434**, 115 (1998), arXiv:hep-ph/9804241.

[209] G. S. Bali, Phys. Rev. **D62**, 114503 (2000), arXiv:hep-lat/0006022.

[210] U. Wolff, Comput. Phys. Commun. **156**, 143 (2004), arXiv:hep-lat/0306017.

[211] A. Pineda, PoS **LAT2005**, 227 (2006), arXiv:hep-lat/0509022.

[212] A. Pineda, J. Phys. **G29**, 371 (2003), arXiv:hep-ph/0208031.

[213] G. S. Bali and A. Pineda, Phys. Rev. **D69**, 094001 (2004), arXiv:hep-ph/0310130.

[214] P. Boyle and G. S. Bali, Nucl. Phys. Proc. Suppl. **106**, 811 (2002), arXiv:hep-lat/0110118.

[215] G. Martinelli and C. T. Sachrajda, Nucl. Phys. **B559**, 429 (1999), arXiv:hep-lat/9812001.

[216] F. Di Renzo and L. Scorzato, JHEP **0102**, 020 (2001), arXiv:hep-lat/0012011.

[217] J. Ambjørn, P. Olesen, and C. Peterson, Nucl. Phys. **B240**, 189 (1984).

[218] C. Anzai, Y. Kiyo and Y. Sumino, Nucl. Phys. **B838**, 28 (2010), arXiv:1004.1562 [hep-ph].

[219] S. Necco and R. Sommer, Nucl. Phys. **B622**, 328 (2002), arXiv:hep-lat/0108008.

[220] J. Engels, F. Karsch, H. Satz and I. Montvay, Nucl. Phys. **B205**, 545 (1982).

[221] V. K. Mitrjushkin, A. M. Zadorozhnyi and G. M. Zinovjev, Phys. Lett. **B215**, 371 (1988).

[222] H. Leutwyler, in *QCD. 20 years later*, edited by P. M. Kastrup and H. A. Zerwas, Aachen 1992 (World Scientific, Singapore, New Jersey, 1993).

[223] G. Boyd, J. Engels, F. Karsch, E. Laermann, C. Legeland, M. Lütgemeier and B. Petersson, Nucl. Phys. **B469**, 419 (1996), arXiv:hep-lat/9602007.

[224] V. M. Braun, in *'95, QCD and high energy hadronic interactions*, edited by J. Tran Thanh Van, Moriond 1995 (Editions Frontières, Gif-sur-Yvette, France, 1995), arXiv:hep-ph/9505317.

[225] M. A. Shifman, in *QCD. 20 years later*, edited by P. M. Kastrup and H. A. Zerwas, Aachen 1992 (World Scientific, Singapore, New Jersey, 1993), arXiv:hep-ph/9304253.

[226] S. Capitani, M. Lüscher, R. Sommer and H. Wittig (ALPHA Collaboration), Nucl. Phys. **B544**, 669 (1999), arXiv:hep-lat/9810063.

[227] W. A. Bardeen, A. J. Buras, D. W. Duke and T. Muta, Phys. Rev. **D18**, 3998 (1978).

[228] O. V. Tarasov, A. A. Vladimirov and A. Y. Zharkov, Phys. Lett. **B93**, 429 (1980).

[229] M. Lüscher and P. Weisz, Phys. Lett. **B349**, 165 (1995), arXiv:hep-lat/9502001.

[230] M. Lüscher and P. Weisz, Nucl. Phys. **B452**, 234 (1995), arXiv:hep-lat/9505011.

[231] A. Bode and H. Panagopoulos, Nucl. Phys. **B625**, 198 (2002), arXiv:hep-lat/0110211.

Acknowledgements

First of all, I am indebted to my parents. I have been able to count on their unconditional support in all these years.

I thank my supervisor Prof. Dr. Gunnar S. Bali for his continuous guidance and help throughout my Ph.D. thesis. I benefited from his experience and physical intuition. He enabled me to go to several conferences and facilitated the long and short-term stays in Barcelona which were vital for this work.

I am truly grateful to Dr. Antonio Pineda for the very fruitful collaboration. I will always remember our intense sessions when neither a university closure over the holidays, nor a general strike could stop us from doing physics.

Special thanks go out to Dr. Christian Torrero. He took me by the hand during my first steps in this field when I started my diploma thesis, and ever since then I have profited from his invaluable help in both physics- and software-related issues.

I thank my colleagues Melanie Pfeuffer, Matthias Aicher, Florian Gruber and Simon Heybrock (by order of appearance). I have benefited from them many times inside and outside physics. I am particularly happy that, in the time we spent together, we talked a lot less about general relativity than about the relativity of physics in general.

I received help from my brother regarding software and other things. I am glad that I can always rely on his assistance.

For my studies I received financial support from the “Bayerische Begabtenförderung” and the “Studienstiftung des deutschen Volkes”. The latter also funded my Ph.D. thesis, together with the “Daimler und Benz Stiftung”. I am grateful to all these institutions.

Thank you, Lena, thank you so much for everything.